



Sistemas Informáticos  
Curso 2004-2005

---

# **EXPERIMENTOS SOBRE LA EVOLUCIÓN DEL LENGUAJE**

Silvia Martín Martín  
Victoria Sanz Trujillo  
Eva Zamorano Valtueña

Dirigido por:  
Prof. Lourdes Araujo Serna  
Dpto. Sistemas Informáticos y Programación (SIP)

---

Facultad de Informática  
Universidad Complutense de  
Madrid

# ÍNDICE:

1. Resumen	2
2. Objetivos	3
2.1. Antecedentes en el estudios de la evolución del lenguaje	3
2.2. Aspectos a estudiar	8
3. Introducción a los Algoritmos Evolutivos	10
4. Sistemas Innatos de Señales	23
4.1. Antecedentes Sistemas Innatos	23
4.2. Descripción de los Sistemas Innatos de Señales	28
4.3. Descripción del Algoritmo Evolutivo para Innatos	30
4.4. Resultados	40
4.5. Conclusiones a partir de los resultados de Innatos	46
5. Experimento de Aprendizaje del Lenguaje	47
5.1. Antecedentes Aprendizaje	47
5.2. Descripción Aprendizaje	52
5.3. Descripción del Algoritmo Evolutivo de Aprendizaje	54
5.4. Primer Experimento de Aprendizaje	62
5.4.1. Resultados del Primer Experimento de Aprendizaje	64
5.4.2. Conclusiones del Primer Experimento de Aprendizaje	67
5.5. Segundo Experimento de Aprendizaje	69
5.5.1. Resultados del Segundo Experimento de Aprendizaje	70
5.5.2. Conclusiones del Segundo Experimento de Aprendizaje	73
5.6. Tercer Experimento de Aprendizaje	74
5.6.1. Resultados del Tercer Experimento de Aprendizaje	76
5.6.2. Conclusiones del Tercer Experimento de Aprendizaje	79
6. Sistemas de Aprendizaje y Selección Natural	81
6.1. Descripción de los Sistemas de Aprendizaje y Selección Natural	81
6.2. Descripción del Algoritmo Evolutivo	83
6.3. Resultados	93
6.4. Conclusiones de Aprendizaje y Selección Natural	97
7. Implementación y Manual de Usuario	99
7.1. Implementación Sistema Innato de Señales	99
7.2. Manual de Usuario Sistema Innato de Señales	101
7.3. Implementación Experimento de Aprendizaje del Lenguaje	106
7.4. Manual de Usuario Experimento de Aprendizaje del Lenguaje	109
7.5. Implementación Aprendizaje y Selección Natural	113
7.6. Manual de Usuario Aprendizaje y Selección Natural	116
8. Palabras Clave	121
9. Referencias	122

# **1. RESUMEN**

## **RESUMEN**

En este proyecto implementamos modelos sencillos de evolución de la comunicación en una población para estudiar distintos aspectos relacionados con la evolución del lenguaje y su ventaja evolutiva.

Consideramos algunos trabajos que han aparecido sobre el tema, implementando el modelo propuesto con un algoritmo evolutivo.

A través de la simulación computacional tratamos de emular sistemas de conocimiento innato, sistemas de aprendizaje y sistemas de aprendizaje con selección natural para el estudio de la evolución real del lenguaje.

## **ABSTRACT**

In this project we implements simple models of evolution of the communication in a population to study different aspects related to the evolution of the language and the evolutionary advantage.

We consider some of the works that have been appeared related to the subject, implementing the proposed model by means of an evolutive algorithm.

Through the computational simulation we try to emulate systems of innate knowledge, learning systems and learning systems with natural selection for the study of the real evolution of the language.

## **2. OBJETIVOS**

### **2.1. ANTECEDENTES EN EL ESTUDIO DE LA EVOLUCIÓN DEL LENGUAJE**

El estudio sobre la evolución del lenguaje es uno de los temas más discutidos dentro de la biología evolutiva.

Todos los animales se comunican entre sí. Podemos definir el lenguaje como una herramienta de comunicación de tal manera que cada ruido o sonido que se emiten los unos a los otros lleven consigo una serie de reglas para poder diferenciar lo que se quiere expresar.

Vamos a comentar algunas de las teorías más importantes en esta evolución:

#### **DARWIN**

Según Darwin [15] nuestro lenguaje ha sufrido múltiples evoluciones, desde, en teoría, un lenguaje basado en ciertos elementos básicos de comunicación como gritos o sonidos, parecidos a la comunicación entre los monos actuales, hasta llegar a un punto tal que los humanos desarrollamos una cierta capacidad para poder crear ciertas estructuras sintácticas y así ir creando un lenguaje de comunicación más complejo.

#### **CALVIN Y BICKERTON**

Bickerton [14] sostiene que el lenguaje actual ha sido desarrollado en dos etapas, en la primera al igual que Darwin dicen que el lenguaje se basó en términos muy básicos de comunicación (un protolenguaje), estos términos, en principio tan sólo nombres y verbos, servían para permitir comunicarse entre los individuos de manera que se pudiesen alertar de ciertos peligros o necesidades básicas como conseguir alimentos. Calvin [14] piensa que nuestros antepasados dispondrían de códigos en el cerebro con los que representa los conceptos que expresan las palabras; se asocian a ciertas palabras ciertas características que vendrían sobreentendidas en dicha palabra, como por ejemplo tigre, indicaría peligro, color amarillo, etc.

En esta fase el lenguaje no tenía ningún tipo de sintaxis. Esta nueva gramática basada en ciertos aspectos sintácticos pudo surgir según Bickerton por las necesidades altruistas dentro de los grupos para crear dentro de estos unas ciertas clases sociales.

### **PINKER Y BLOOM [14]**

Defienden la hipótesis mediante la cuál nuestra capacidad lingüística se ha podido producir mediante una selección natural. Esta selección partió inicialmente de un esquema básico y este esquema fue evolucionando hasta llegar a lo que tenemos hoy día. Según ellos esta evolución no ha podido surgir de la nada y ha tenido que originarse con alguna función específica, aquí entra dicha selección natural ya que sin ella no se podría haber creado una estructura tan compleja. Para intentar explicar esta teoría se están realizando ciertos estudios en varios campos entre los que entran la paleontología para intentar saber más sobre la estructura biológica de ciertos especímenes y así conocer algo sobre como pudieron llegar a desarrollar el lenguaje mediante estudios del aparato fonatorio. Otro de los campos en los que se están realizando continuos avances para poder explicar como ha podido surgir esta selección del lenguaje es en el campo de la genética que tanto ha descubierto últimamente, en este terreno han llegado a descubrirse ciertos genes que podían llegar a ser los responsables de la posible evolución del lenguaje, así como otros producen un estancamiento de dicha evolución.

### **DEACON**

Deacon [16] defiende el termino de comunicación simbólica, es decir, que nuestro lenguaje ha surgido a partir de una creación de ciertos símbolos. Deacon piensa que el lenguaje surge de una cierta necesidad para comunicar las experiencias de cada individuo. Esta comunicación se iba transmitiendo de generación en generación y esta evolución ha venido originada también por los cambios producidos en nuestra anatomía.

### **TORO Y CASTRO [14]**

Según ambos el lenguaje ha sido originado a partir de un lenguaje básico, es decir se basan en las teorías de Darwin. Este lenguaje evoluciono mediante una selección natural. Esta selección supuestamente ha sido producida como consecuencia

de la creación de una clasificación de conductas dentro de las distintas tribus. Estas conductas podrían haberse clasificado en buenas y malas, en estas tribus los individuos más dotados serían los encargados de llevar a cabo este ‘supuesto’ proceso de selección. Estos aceptaban el aprendizaje a las generaciones posteriores de las conductas que ellos clasificaban como aceptables, mientras que intentaban eliminar de la conducta aquellas que no deberían realizarse, a través de esta necesidad de comunicación entre los individuos pudieron surgir las primeras evoluciones del lenguaje.

## **SIMON KIRBY**

Kirby [2] trata de explicar por qué el lenguaje natural es un entorno adecuado para aplicar métodos de vida artificial. Analiza algunas de las características que hacen al lenguaje el sistema natural más complejo del que tenemos conocimiento, y lo hace a través de una aproximación cronológica: desde la evolución de sistemas de señalización simples innatos, la evolución cultural de sistemas de señales no innatos sino aprendidos, los fundamentos de las señales y el modelado de espacios flexibles significativos, el surgimiento de complejos y estructurados lenguajes a través de la “negociación” y de la evolución cultural y de modelos que abordan la evolución de los lenguajes y los mecanismos de aprendizaje de dichos lenguajes.

La vida artificial proporciona la respuesta al problema metodológico de unir teorías y predicciones para complejos sistemas dinámicos en los que interactúan muchos componentes. Los modelos microsimulados, diseñados cuidadosamente, proporcionan una herramienta para probar ideas y teorías acerca de cómo se producen fenómenos naturales como las bandadas de pájaros o la división celular. Desde el punto de vista de la vida artificial, el lenguaje humano es un campo ideal para profundizar.

## **MICHAEL OLIPHANT**

La mayoría de las comunicaciones que percibimos son el resultado de comportamientos puramente innatos. Oliphant [5] sugiere la utilidad de estudiar la comunicación como una combinación de explotación y manipulación. Define

comunicación como “una cadena casual de sucesos en los que un individuo, emisor muestra un comportamiento en respuesta a una situación en particular, y un segundo individuo, el receptor, responde a este comportamiento. Esta interacción es comunicativa si involucra la manipulación por parte del emisor y la explotación por parte del receptor”.

Nos muestra ejemplos de comunicación animal como por ejemplo el de los monos, que se comunican a base de la producción de ciertos sonidos para identificar cada tipo de peligro, así cuando aparece un animal del tipo de leopardo el sonido que producen suele ser más suave que para alertar de que algún ave se encuentra cerca, de tal manera que si producen este último tipo de sonido tienen identificado también que el daño proviene del cielo.

Este comportamiento que tienen los monos se podría definir como una simple definición de comunicación, un mono (emisor) realiza un cierto sonido (mensaje) que es recibido por otro individuo, en este caso el receptor que tiene una reacción en función del mensaje que haya recibido.

Este tipo de comunicación se suelen denominar interacciones. Estas interacciones se pueden clasificar de una forma un tanto peculiar, según estas produjesen una manipulación o una explotación en función de emisor y el receptor, en caso en el que el emisor obtuviese una beneficio en dicho envío esta interacción se denomina manipulación mientras que si el beneficio es para el receptor la interacción se denomina explotación, tan sólo proporciona la definición de comunicación en el caso en el que el beneficio entre el emisor y el receptor fuese mutuo, este beneficio se podría traducir como una respuesta al mensaje producido por alguno de los implicados.

Manipulación (Beneficio del receptor):

$$(X \rightarrow Y) \Longleftarrow (Y \rightarrow Z)$$

Explotación (Beneficio del emisor):

$$(X \rightarrow Y) \implies (Y \rightarrow Z)$$

Comunicación:

$$(X \rightarrow Y) \rightleftarrows (Y \rightarrow Z)$$

Es en las teorías de los dos últimos autores en el que hemos basado mayormente nuestro trabajo.



## **2.2. ASPECTOS A ESTUDIAR**

La evolución del lenguaje es uno de los problemas más importantes investigados en el campo de la Vida Artificial. El lenguaje natural es quizás el sistema natural más complejo que existe.

Tanto para la lingüística como para la evolución, es importante comprender los orígenes de este comportamiento único y derivar las propiedades de éste a partir de los sistemas no lingüísticos.

Pues bien, el objetivo de nuestro trabajo es implementar modelos sencillos de evolución de la comunicación en una población, con el fin de estudiar los distintos aspectos relacionados con la evolución del lenguaje y su ventaja evolutiva.

En el mundo científico han aparecido trabajos relacionados con este estudio. Nosotras hemos considerado algunos de ellos, implementando el modelo propuesto con un algoritmo evolutivo.

Consideraremos algunos modelos representativos de los siguientes aspectos del problema:

### **Evolución biológica de sistemas innatos de señales simples:**

La mayoría de las comunicaciones elementales están guiadas en buena parte por capacidades y tendencias innatas. Las funciones de transmisión y recepción de información vienen dadas por el genoma del agente.

La capacidad lingüística ha evolucionado como resultado de la selección natural desde los primates, capaces de emitir gritos inarticulados, hasta los seres humanos dotados de lenguajes articulados y reglas sintácticas.

Estudiaremos en qué condiciones y con qué características surgen este tipo de sistemas.

### **Evolución de sistemas de señales aprendidos:**

En las lenguas humanas, la correspondencia entre las señales y los significados es fundamentalmente aprendida. Una característica interesante es que la introducción de aprendizaje lleva a que surja otro sistema complejo: la cultura. En este tipo de sistemas se produce un proceso “histórico” de transmisión de la información mediante comportamientos repetitivos de aprendizaje y puesta en práctica. Un agente aprendiz “más joven” (es decir, ha sido miembro activo de la población durante menos tiempo) realiza preguntas para la adquisición de conocimiento a individuos adultos. La comunicación entre ambos hará que el conocimiento de los aprendices se amplíe de manera que evolucione su lenguaje.

### **Evolución de sistemas de señales innatos y aprendidos:**

Este es el sistema que más se aproxima al sistema real. Mezcla los conocimientos adquiridos con comportamientos puramente innatos con los obtenidos mediante aprendizaje.

Durante una parte de su vida el individuo aprendiz va a ser capaz de adquirir conocimientos de la población adulta a base de preguntar sobre su conocimiento. Este aprendizaje se realizará en varios ciclos hasta que alcance el conocimiento correspondiente para pasar a ser adulto o hasta que cumpla la edad límite de vida en la que este individuo morirá, ya sea aprendiz o adulto.

### **3. INTRODUCCIÓN A ALGORITMOS EVOLUTIVOS**

Los Algoritmos Evolutivos se utilizan cada vez más como una herramienta útil al enfrentarnos a problemas de ingeniería. Es una rama de la computación que imita la evolución de la naturaleza ayudándose de la biología.

Con los AE's podemos implementar por una parte sistemas capaces de adaptarse dinámicamente a un entorno cambiante y por otra parte diseñar sistemas artificiales que simulen la evolución natural *para solucionar problemas de optimización arbitrarios*.

El esquema básico del algoritmo sería el siguiente:

- 1.- Se crea una población inicial de individuos que representan, de distintas formas en función del tipo de algoritmo evolutivo elegido, soluciones potenciales al problema. Cada individuo contiene una cadena de genes, que en los algoritmos genéticos será una cadena de bits.
- 2.- Definimos una función de evaluación que decidirá cuáles son los mejores individuos respecto a sus posibilidades de ser solución al problema tratado.
- 3.- A partir de la población inicial, se realiza la reproducción dando mayor probabilidad de reproducirse a aquellos individuos con mejores resultados con respecto a la función de evaluación.
- 4.- Volvemos a evaluar la población dando lugar a las siguientes generaciones con los mejores individuos obtenidos a partir de una selección realizada entre los individuos de la población.
- 5.- Con el paso de las iteraciones o generaciones la población tenderá a ser mejor ya que la función de evaluación favorece a la supervivencia de los mejores individuos frente a los peores.
- 6.- Después de un determinado número de generaciones la población irá convergiendo dando lugar a la solución óptima del problema en la mayoría de las situaciones.

A partir de 1970 los algoritmos evolutivos se especializan en diferentes técnicas evolutivas, como son los algoritmos genéticos, la programación evolutiva, sistemas de aprendizaje y programación genética, entre otros. Todos estos métodos comparten cuatro características principales: población inicial de individuos, función de evaluación, capacidad de crear y matar individuos según van pasando las generaciones y por último la capacidad de cada individuo de heredar aptitudes de sus progenitores.

Vamos a estudiar detalladamente cada uno de estos métodos:

### **ALGORITMOS GENÉTICOS:**

En los años setenta John Holland propuso los algoritmos genéticos inspirados en la evolución biológica y en la genética.

Estos algoritmos parten de una población inicial, donde cada individuo se representa por una cadena de bits, que evoluciona de acuerdo a una serie de acciones, como la evaluación, selección, cruce, mutación...hasta que se llegue al número de generaciones fijadas de antemano o bien se llegue a un punto en el que la población no sufre cambios significativos.

Se muestra a continuación el pseudo-código que define el esquema básico de un AG simple:

*BEGIN*

*Crear población inicial;*

*Evaluar la población y asignarle una puntuación a cada individuo dependiendo de sus aptitudes;*

*WHILE NOT Terminado DO*

*BEGIN*

*Seleccionar los padres de entre los individuos de la población favoreciendo a aquellos que tengan una puntuación superior.*

*Cruzar los genes de los padres dando lugar a los genes de los hijos*

*Mutar los hijos*

*Evaluar de nuevo la población y elegir los candidatos para la siguiente generación*

*IF número de generaciones = a las indicadas o bien se ha encontrado la solución deseada THEN*

*Terminado:= TRUE;*

*END;*

*END.*

Ahora comentaremos cada uno de los pasos para la implementación de los AG:

### **Codificación del genoma:**

Antes de comenzar el desarrollo del AG propiamente dicho es necesario que realicemos una codificación de los genes que representan a cada individuo. La decisión de cómo definir el lenguaje no es sencilla ya que de ella depende el encontrar o no una solución óptima.

Una de las codificaciones más utilizadas para representar el genoma de los AG es el código de Gray, que establece que los genes de dos individuos vecinos sólo difieren en un bit.

### **Generación de la población inicial:**

La población inicial suele crearse de forma aleatoria, es decir, los genes que contenga cada individuo no se fijan de antemano sino que se deja que la evolución actúe sobre ellos.

El tamaño elegido para la población inicial debe ser lo suficientemente grande para garantizar la diversidad de las soluciones. Así cuanto mayor sea la población más exhaustivo será el estudio del problema. Pero por otro lado, si el tamaño es demasiado grande, el tiempo necesario para su evolución también será grande. Por supuesto, depende del ordenador en el que se esté ejecutando.

Por eso debemos llegar a un compromiso entre ambos, para crear un estudio donde los resultados sean fiables, pero sin desperdiciar recursos.

### **Función de evaluación:**

La elección de la función de evaluación depende exclusivamente del problema que se esté tratando. Dicha función debe plasmar de forma clara las características que ha de tener la población que estamos estudiando, ya que a partir de ella, nuestro programa tendrá que elegir a los mejores individuos. Éstos serán los que mejor puntuación obtengan en dichas características.

### **Reproducción o Cruce:**

Antes de realizar la reproducción mediante la operación de cruce es necesario definir qué y cuántos individuos van a reproducirse y cuántos descendientes van a dar lugar. Para ello definimos una función de selección que será la encargada de decidir que individuos van a ser los participantes en el cruce. Las dos formas más habituales de implementar esta función son con una probabilidad de cruce fija o bien con una probabilidad dada por la función de evaluación, favoreciendo con esta última a los mejores individuos de la población.

Una vez seleccionadas las parejas que van a reproducirse, se realiza el cruce, que consiste en la recombinación de las cadenas de genes de los progenitores dando lugar a las nuevas cadenas de los descendientes, adquiriendo éstos las propiedades de los padres.

Los tipos de cruce más usuales son el monopunto, que elige un bit a partir del cual se intercambian los genes de los padres. El cruce doble, en el que elegidos dos bits se intercambian todos los genes que hay entre ellos. El cruce con máscara, en el que se genera aleatoriamente un vector que actuará de máscara y con el que se filtrarán las cadenas de genes de los padres dando lugar a la de los hijos.

El cruce se encarga de aprovechar y propagar las mejores particularidades que definan la población realizando una búsqueda en profundidad.

## **Mutación:**

La mutación consiste en una pequeña alteración de las cadenas de genes de los individuos generados tras el cruce. Esta viene dada por un factor de mutación que define la probabilidad de mutar o no cada uno de los bits que componen la cadena de genes del individuo. Esta probabilidad es bastante menor que la probabilidad de cruce ya que de lo contrario no se llegaría a la convergencia esperada de la población.

La mutación realiza una búsqueda en anchura, permitiendo así analizar nuevos dominios del espacio de búsqueda, intentando así obtener mejores soluciones.

## **¿Quién sobrevive?**

Una vez que tenemos las dos poblaciones, la de los progenitores y la de descendientes, debemos decidir cuáles son los individuos que pasarán a la siguiente generación. Esta nueva población va a depender del objetivo de nuestro programa. La solución más sencilla sería elegir a los mejores individuos. Así llegaríamos antes a una solución del problema pero exploraría menos espacios de soluciones, con lo que podríamos devolver como solución un óptimo local, no siendo ésta la mejor solución para un problema de optimización.

En general existen varios criterios para el reemplazo de los individuos:

1. Reemplazo inmediato: Los descendientes sustituyen directamente a sus progenitores.
2. Reemplazo con factor de llenado: Los descendientes sustituyen a los progenitores que más se les parezcan, no siendo necesariamente sus padres.
3. Reemplazo por inclusión: Se juntan las dos poblaciones en una sola, de la cual solo pasará a la siguiente generación los  $N$  individuos mejores siendo  $N$  el tamaño de la población de progenitores.
4. Reemplazo por inserción :

Si el tamaño de la población de progenitores es mayor o igual que la de descendientes, se eliminan los  $S$  individuos peores de la primera población (siendo  $S$  el tamaño de la población de descendientes), y se sustituyen por los hijos.

Si el tamaño de la población de progenitores es menor que la de descendientes, se eligen los N mejores descendientes y pasan directamente a la siguiente generación.

### **PROGRAMACIÓN EVOLUTIVA:**

Desarrollada por Lawrence J. Fogel en 1960. La programación evolutiva se aplica principalmente a problemas de optimización y a problemas de aprendizaje mediante autómatas finitos. A partir de la programación evolutiva obtenemos sistemas de aprendizajes basados en la evolución mediante la selección natural.

Es un algoritmo fácil de implementar muy similar al de los algoritmos genéticos pero en el que la generación de nuevos individuos se realiza simplemente mutando los individuos de la población actual, al no haber cruce entre los individuos, no existe relación entre las especies. Lo habitual es que el tamaño de la población sea el tamaño N de la población inicial.

Un esquema básico que define dicho algoritmo sería el que sigue:

1. Generación aleatoria de individuos que formarán una población inicial.
2. Duplicación de población actual que pasará a ser mutada
3. Aplicamos la mutación sobre los individuos duplicados. El factor de mutación no tiene por qué ser constante en la generación de individuos.
4. Selección mediante torneo de los individuos. Comparando los individuos antiguos con los mutados y seleccionamos los que pasarán a la siguiente generación según nuestra función objetivo.

Realizamos el cálculo de la ganancia para cada individuo, de manera que dado un individuo X de la población conjunta (padres e hijos), se seleccionan aleatoriamente otros S individuos y la ganancia para cada uno de ellos es el número de estos individuos que tienen peor solución con respecto a la objetivo que X. Hay dos formas de realizar la selección:

- Se eligen los N individuos de la población total (padres y mutados) que mayor ganancia tengan
- Se eligen los N individuos mejores de la población total (padres y mutados) con respecto a la función de evaluación



No hay forma preestablecida para la representación de los individuos en la programación evolutiva. Se suelen utilizar autómatas finitos para representar los individuos en el que los genes de estos representen a los distintos estados del autómata.

### **ESTRATEGIAS EVOLUTIVAS:**

Las estrategias evolutivas son parte de los algoritmos evolutivos. Fueron desarrolladas en 1964 en Alemania por Ingo Rechenberg, Hans-Paul Schwefel y Paul Bienert.

Fueron creadas como una técnica experimental de optimización que trabajaba sobre poblaciones de individuos del dominio de los números reales.

Como características generales de esta técnica se pueden destacar su eficiencia, su buena disposición para optimizar valores reales, su gran desarrollo teórico y la autoadaptación de los parámetros.

La representación utilizada por esta estrategia son vectores de valores reales, la selección de los padres es aleatoria, su recombinación es discreta y la mutación gaussiana.

De manera general, independientemente de la variación de la estrategia utilizada, cada individuo de la población es un posible óptimo de la función objetivo. Cada individuo está representado por dos tipos de variables: las objeto que son los valores posibles que hacen que la función objetivo alcance el óptimo global y las estratégicas que indican de qué manera la mutación afecta a las variables objeto.

La primera versión de esta técnica fue la EE-(1+1), llamada así porque en cada generación sólo había dos individuos, un padre y un hijo. Siendo el mejor de ambos el que pasaba a la siguiente generación.

El hijo se obtenía aplicando la mutación sobre el progenitor, dicha mutación estaba asociada a una distribución normal de media 0 y varianza  $\sigma^2$ .

Rechenberg desarrolló una teoría sobre la velocidad de convergencia llamada regla de 1/5-éxito.

La mutación se adapta respecto a ésta regla de la siguiente manera:

$$\sigma' = \text{mut}(\sigma) = \sigma/c \text{ si } p > 1/5$$

$$\sigma * c \text{ si } p < 1/5$$

$$\sigma \quad \text{si } p = 1/5$$

Siendo  $c=0,85$  un posible valor.

La forma de actuar de ésta regla es la siguiente:

Cada vez que se hayan realizado en mutaciones se cuentan los éxitos que ha habido en las  $10 \cdot n$  mutaciones anteriores. Siendo  $p$  el porcentaje de éxitos, si éste es mayor que  $1/5$ , se multiplica la longitud de paso por  $c$ , y si es menor que  $1/5$ , se divide entre  $c$ .

La estrategia del 1/5-éxito puede hacer converger el algoritmo demasiado rápido, ya que cuando no se producen éxitos la longitud del salto se vuelve muy pequeño. Así que la mayoría de las veces ésta técnica se quedara con óptimos locales.

Debido a estos inconvenientes con la regla 1/5-éxito, se continuó investigando hasta encontrar otras variantes de la estrategia evolutiva, que fueron  $EE-(\mu, \lambda)$  y  $EE-(\mu + \lambda)$  con  $\mu > 1$  padres y  $\lambda > 1$  descendientes, cuya característica más importante es la auto-adaptación de los parámetros. Este proceso se basa en una relación indirecta entre dichos parámetros y el valor que toma la función objetivo.

La única diferencia entre  $EE-(\mu, \lambda)$  y  $EE-(\mu + \lambda)$  es que en el último se selecciona la población que pasa a la siguiente generación de entre los  $\mu + \lambda$  individuos que resultan de unir la población de padres e hijos, tomando de ellos los  $\mu$  mejores.

Esta variante de estrategia evolutiva destaca la búsqueda local antes que la global. Pero a la vez pierde efectividad en su auto-adaptación.

## **PROGRAMACIÓN GENÉTICA:**

La programación genética nació a principios de los ochenta y contó con John Koza como su máximo representante.

La gran diferencia de la programación genética con el resto de algoritmos evolutivos es que ésta en lugar de representar soluciones con cadenas de bits, utiliza programas ejecutables.

Para la representación de dichos programas se utilizan estructuras no lineales como árboles o grafos. En estos árboles hay nodos que serán llamadas a funciones y nodos que representarán los argumentos correspondientes a las funciones, que serán los hijos de los primeros. Además, el tamaño de esta estructura no tiene que ser fijo, sino que puede variar tanto en anchura como en profundidad.

La programación genética sigue un proceso parecido al de los algoritmos genéticos. Lo primero que se debe hacer es concretar las funciones y símbolos terminales que definirán el programa. Estos han de ser suficientes para resolver el problema. Y cada función debe ser capaz de aceptar como argumento cualquier símbolo terminal de los definidos y además cualquier valor devuelto por otra función.

Tras esto ya se puede crear una población inicial de árboles mediante una composición de funciones y símbolos terminales, la mayoría de las veces ésta inicialización se realiza de forma aleatoria.

A partir de aquí se repetirán los pasos hasta alcanzar la condición de terminación del algoritmo.

Se debe ejecutar cada programa y asignarles una puntuación que obtendremos con la función de aptitud. Esta función mide el acierto de los individuos al ser ejecutados programas que ya sabemos qué valores deben alcanzar.

El siguiente paso a realizar es la reproducción, donde lo primero que se debe hacer es la selección de los puntos de corte de los padres que se hayan decidido en la función de selección, que casi siempre se basa en la función de evaluación.

Para conseguir un algoritmo más eficiente debemos intentar dar menor probabilidad de selección a los arcos que llevan a los nodos hojas o terminales. De esta manera conseguimos más desigualdad entre los padres y los hijos.

Una vez que hayamos elegido de forma aleatoria un punto de cruce en el primer árbol, debemos tener en cuenta ciertas limitaciones para la elección del punto de cruce

del segundo árbol. Dichas limitaciones pueden ser la concordancia de tipos de los subárboles a cruzar y que el tamaño sea oportuno, ya que mediante el cruce se hace que los árboles tiendan a crecer mucho si no limitamos el tamaño máximo que pueden alcanzar. Por lo tanto, habitualmente, si el segundo subárbol elegido para el cruce induce a que el hijo alcance un tamaño superior al que consideramos máximo, es mejor no realizar dicho cruce y probar de nuevo con otro. Aun así, es habitual que tras el cruce el tamaño de los hijos sea mayor que el tamaño de los padres.

Una vez seleccionados los puntos de cruce de los dos árboles a reproducir, lo único que hay que hacer es intercambiar los subárboles que cuelgan de los nodos elegidos.

El cruce, en programación genética, de dos individuos parecidos no genera hijos iguales, como ocurría en los algoritmos genéticos. Podemos cruzar dos árboles iguales con una buena puntuación y que sin embargo los descendientes no sean convenientes para encontrar la mejor solución. Y es que con la programación genética el espacio de búsqueda es muy grande. Por ello es necesario que se utilicen tamaños de población mucho más grandes que los que se utilizan en los algoritmos genéticos.

A continuación lo que se debe hacer es la mutación, y para ello existen varias formas de actuar, unas producirán descendientes más diferentes a los progenitores y otras producirán descendientes prácticamente iguales a sus progenitores. De todas formas, en programación genética, no es tan importante conseguir diversidad mediante la mutación, ya que se consigue con el cruce.

Los métodos más utilizados para realizar la mutación son la mutación en un punto, que elige un nodo y se modifica su valor por otro compatible. Este nodo compatible se elige aleatoriamente de una lista que habremos creado con todos los nodos terminales.

La otra forma es la mutación por subárbol aleatorio, que elige una arista al azar y la sustituye por un subárbol generado de forma aleatoria.

Por último se debe decidir qué individuos pasan a la siguiente generación. Para ello, y como ocurría en los algoritmos genéticos, se utiliza una función de evaluación que determinará la bondad de los programas. Esta función se basa en dar al programa una serie de datos y cotejar los resultados que devuelve el programa con los resultados esperados.

## **PROGRAMAS EVOLUTIVOS**

Surgen en 1994 como una variación de los Algoritmos Genéticos de la mano de Michalewicz y toman el nombre de Programas Evolutivos o Programas de Evolución (PEs). Los primeros PEs fueron realizados para la resolución de problemas de optimización paramétrica de alta precisión.

Aparecieron debido a la rigidez de la representación de los individuos en los Algoritmos Genéticos (cadenas binarias). Los PEs utilizan representaciones más naturales y más cercanas al dominio del problema.

Las estructuras de datos más utilizadas para la representación en los PEs son vectores, matrices, listas, conjuntos, árboles, etc. Y los operadores genéticos son específicos de la representación.

Existen dos grandes grupos de clasificación de los PEs según la naturaleza de los objetos tratados:

- Especializados en optimización numérica
- Especializados en optimización combinatoria

Sobre los PEs de optimización numérica se han realizado numerosos estudios que han dado como resultado la conveniencia de utilizar como representación un vector de números de punto flotante, es decir, cada individuo estará representado por un vector de estas características, en el que cada elemento se corresponde con un gen.

Además el cruce da lugar a dos hijos cuyo vector de representación será el promedio de los valores de los vectores de los padres.

La mayor ventaja que tienen este tipo de PEs es que al utilizar valores en punto flotante para el espacio de búsqueda, la solución será más exacta y como los valores estarán más próximos al dominio del problema será más fácil la incorporación de conocimiento.

Sin embargo, al utilizar valores sin codificar, cada vez que se obtengan resultados habrá que comprobar que se encuentren entre los rangos permitidos. Para ello se introducen restricciones que deberán ser tratadas. En algunos casos al encontrar valores fuera de dichos rangos se podrán eliminar directamente, mientras que en otras ocasiones será más conveniente utilizar mecanismos más complejos para su tratamiento.

Para el caso de los PEs especializados en optimización combinatoria no existe ninguna representación preestablecida y los operadores genéticos serán específicos de cada representación.

Algunas representaciones de ejemplos concretos que utilizan este tipo de Programas Evolutivos son los siguientes:

- El Viajante de Comercio

El problema se basa en que dadas  $m$  ciudades y los costes de viajar de unas a otras se desea calcular el recorrido completo, cerrado y conexo de coste mínimo.

La representación que se utiliza para este problema es una cadena de números enteros debido a su sencillez y eficacia. Dicha cadena de enteros representa el orden en que el viajante ha recorrido las ciudades, por lo tanto esa cadena será lo que represente a cada individuo de la población y serán las posibles soluciones al problema planteado.

Un posible cruce a utilizar es el emparejamiento parcial, que consiste en la elección de dos puntos de corte, intercambiando los elementos que queden en el rango intermedio entre ambos progenitores:

Padre 1 = (1,2,3,4,5,6,7,8,9)

Padre 2 = (9,7,5,3,1,2,4,6,8)

Dando lugar a los siguientes hijos tras el intercambio de los elementos entre los puntos de corte 4 y 6 y tras una serie de transformaciones para evitar conflictos por repetición. Cuando haya dos elementos iguales en un mismo individuo se intercambia el elemento por su pareja en el intercambio:

Hijo 1 = (4,5,6,3,1,2,7,8,9)

Hijo 2 = (9,7,1,4,5,6,3,2,8)

Otros posibles operadores de cruce que se podrían utilizar son el cruce por orden, el cruce por ciclos, el cruce por recombinación de rutas, etc.

- Planificación de Horarios

Este problema trata de una empresa en la que se dispone de  $m$  empleados que deben atender  $k$  puestos durante  $n$  turnos a lo largo de la semana. El objetivo

es conseguir el horario que cumpla las siguientes condiciones: Siempre ha de haber un solo empleado en cada puesto, no pudiendo estar un mismo empleado en dos sitios en el mismo turno y siendo el número de empleados parados el mínimo.

En este ejemplo concreto la representación utilizada es una matriz de enteros, en la que cada posición  $\langle i,j \rangle$  corresponde al puesto que le ha correspondido al empleado  $i$  en el turno  $j$ .

La función de aptitud que se utiliza es la suma de productividades, sufriendo la aptitud penalizaciones debido a las condiciones que tiene que cumplir.

El cruce que utiliza es heurístico de orden  $q$ , es decir, se intercambian los  $q$  mejores entre los progenitores, siendo esos mejores elegidos a partir de la aptitud local asociada al horario asignado.

La mutación será de orden  $p$ , en la que para un solo empleado toman dos sucesiones seguidas de  $p$  turnos y se intercambian.

#### Conclusiones en el diseño de los Programas Evolutivos:

Se ha de elegir una representación para el problema lo más fácil de manejar posible y que se aproxime tanto como se pueda a la naturaleza real del problema.

En la población inicial creada deberá haber, como mínimo, algún individuo factible.

Los operadores genéticos han de ser acordes a la representación seleccionada y manejar de forma correcta las restricciones creadas.

La función de evaluación ordenará las soluciones de acuerdo a su aptitud.

Para la realización del estudio de nuestros sistemas hemos elegido algoritmos evolutivos que se acerquen lo máximo posible a la representación real de una población y su lenguaje. Se trata de los Programas Evolutivos, que utilizan representaciones más naturales y más cercanas al dominio del problema.

## **4. SISTEMAS INNATOS DE SEÑALES**

### **4.1. ANTECEDENTES SISTEMAS INNATOS**

La mayoría de los estudios realizados demuestran que la mayor parte de los sistemas de comunicación son innatos. Todos los seres vivos nacen con la capacidad de enviar y recibir señales. Durante su vida este conocimiento innato se irá modificando mediante selección natural.

La Programación Evolutiva ha sido utilizada para tratar de resolver problemas computacionales mediante técnicas de simulación inspiradas en la evolución natural de los seres vivos.

Hamilton (1964; 1967), Maynard Smith (1982) utilizaron la teoría evolutiva para resolver el problema de decisión en los juegos. Cada vez que un individuo jugara con otros miembros de la población obtendría un beneficio que es utilizado para calcular la aptitud, necesaria para la selección.

Werber y Dyer (1991) simularon un entorno dónde sólo había dos tipos de individuos: hembras y machos. Las hembras de su modelo eran inmóviles pero podían ver, sin embargo los machos se podían mover, pero eran ciegos. En su estudio observaron cómo evolucionaba el sistema de señales que permitían a las hembras guiar a los machos hasta su posición.

Levin (1995) utilizó algoritmos genéticos para estudiar la evolución de la correspondencia entre los estados de un agente interno y su comportamiento externo. Unos buscadores eran guiados por el agente interno para buscar “comida” en un determinado entorno y estudiar como la comunicación puede evolucionar para facilitar la coordinación de comportamientos externos.



Canguelosi y Parsini [1] usaron un modelo de redes neuronales para la clasificación de alimentos. Algunos alimentos son comestibles (que da energía al individuo que la come) y otros alimentos son venenosos (por lo que el individuo pierde energía). La capacidad de diferenciar la comida es la que proporciona la aptitud de cada individuo, lo que condiciona la evolución de la población. Mediante la simulación de la comunicación entre individuos, aquellos individuos que no pueden observar las características de los alimentos pueden aprender a clasificarlos correctamente.

Michael Oliphant [5] realiza una revisión de los trabajos que se han realizado sobre la comunicación innata en la vida artificial y expone un modelo de simulación propio.

Oliphant ve la comunicación como una combinación de Explotación y Manipulación.

En este contexto Explotación se refiere a la situación en la que el comportamiento de un individuo está determinado de alguna manera por la forma en que otro individuo se ha comportado en una situación particular, y ese comportamiento puede ser observado. Por ejemplo, si un empleado que va a pedir un aumento de sueldo al director oye al director discutir con otro compañero, puede decidir no pedir el aumento de sueldo en ese momento ya que ha observado el comportamiento colérico del director.

Manipulación es justo lo contrario de Explotación. Se refiere a la situación en la que el comportamiento de un individuo está determinado, de alguna manera, por la forma en que se espera que otro individuo pueda responder a ese comportamiento. Por ejemplo, si una persona grita “Socorro” está manipulando el hecho de que la persona que lo oiga acuda en su ayuda.

En el modelo de Oliphant cada individuo conoce una serie de pares señal-significado. Para la comunicación con otros individuos tiene una función de transmisión  $s$  y una función de recepción  $r$ . Para un significado  $\mu$ , la función  $r(\sigma, \mu)$  será la probabilidad de que la señal  $\sigma$  ha sido interpretada correctamente por el receptor.

Mediante estas funciones puede calcular la probabilidad de que las señales mandadas por la función  $s$  sean interpretadas correctamente por la función de recepción  $r$ . La probabilidad de acierto la calcula mediante la función:

$$ca(s, r) = \frac{1}{|M|} \sum_{\mu} \sum_{\sigma} s(\mu, \sigma) r(\sigma, \mu)$$

donde  $|M|$  es el número de significados.

S y R se definen como las funciones de probabilidad media para una población de  $n$  agentes:

$$S(\mu, \sigma) = \frac{1}{N} \sum_i s_i(\mu, \sigma) \quad R(\sigma, \mu) = \frac{1}{N} \sum_i r_i(\sigma, \mu)$$

La probabilidad tomará valores entre 0.0 y 1.0, siendo el significado del valor 1.0 que todas las comunicaciones de todos los individuos han sido acertadas.

Oliphant establece que para realizar la simulación hay que describir el marco de trabajo de la simulación y los parámetros de la simulación.

Para definir el marco de trabajo es necesario primero elegir una representación adecuada del genoma de los individuos. Un ejemplo de la estructura del genoma para el modelo de Oliphant, donde se especifican las funciones de envío y recepción, sería la siguiente:

Genes de transmisión			Genes de recepción		
Señal realizada			Significado interpretado		
Significado a codificar	0	d	Señal observada	a	1
	1	a		b	2
	2	b		c	3
	3	c		d	0

Para esta estructura una posible representación genética del individuo sería:

d	a	b	c	1	2	3	0
---	---	---	---	---	---	---	---

Una vez establecida la representación genética se tienen que definir funciones para poder evaluar la aptitud de cada individuo. En este modelo la evaluación de un individuo se realiza mediante la comunicación con otros individuos seleccionados aleatoriamente. Cuando se produce un acierto en la comunicación, se pueden beneficiar el receptor o el receptor y el emisor, según los parámetros de simulación que se estén utilizando.

Para calcular la aptitud de los individuos en cada iteración, se divide el número de aciertos en la comunicación por el número total de comunicaciones. El valor de la aptitud de cada individuo determina el número de descendientes que tendrá, ya que indica la probabilidad de ser seleccionado para la reproducción.

En su experimento Oliphant intenta encontrar los parámetros de simulación bajo los cuales la función  $ca(S,R)$  alcanzará el valor 1.0.

Uno de los parámetros que estudia para ver como evoluciona el lenguaje, es la forma en que se calcula la aptitud en función de la respuesta de éxito o fallo en la comunicación. Considera dos alternativas:

- Beneficio mutuo: el emisor y el receptor se benefician del éxito en la comunicación.
- Comportamiento altruista: solo se beneficia el receptor del éxito de la comunicación.

Observa que los mejores resultados los obtiene mediante beneficio mutuo, mientras que con el comportamiento altruista la simulación evolutiva no garantiza el éxito.

Otro de los parámetros que estudia es la manera de seleccionar los individuos con los que comunicarse:

- Selección aleatoria: se toman individuos aleatoriamente para la comunicación.
- Selección espacial: la población está organizada espacialmente y se toman parejas cercanas para la comunicación.

Oliphant observa que mediante selección espacial los valores de probabilidad de acierto en la comunicación que alcanzan son casi óptimos, tanto con beneficio mutuo como con el comportamiento altruista. Opina que la organización espacial es equivalente a la selección de parientes (*kin selection*). Si se tiene mayor probabilidad de hablar con los vecinos, es más probable que hablemos con un pariente. Es decir, un individuo genéticamente muy parecido a ti, por lo que el acierto en la comunicación será mejor.

Para nuestro sistema innato de señales utilizaremos una representación parecida a la de Oliphant. También variaremos los mismos parámetros de simulación para ver como evoluciona el lenguaje de nuestro sistema.

## **4.2. DESCRIPCIÓN DE LOS SISTEMAS INNATOS DE SEÑALES**

El sistema que vamos a estudiar simula la evolución de sistemas innatos de señales, basándose para ello en poblaciones cuyos individuos tienen un conocimiento adquirido en el momento del nacimiento, es decir, sólo poseen conocimiento innato, en ningún momento de su vida hay aprendizaje alguno.

Hay diversos factores que vamos a estudiar en este tipo de sistemas. A partir de ellos podremos ver cuáles serán las mejores condiciones para que la población acabe convergiendo a un lenguaje común, o lo que es lo mismo, para que cualquier individuo de la población se comunique con otro y se entiendan.

El primer factor que estudiaremos será la comunicación entre los individuos de la población. Dicha comunicación se basa en el entendimiento o no de mensajes enviados de unos individuos a otros.

Experimentaremos con dos posibles formas de comunicación: la comunicación vecinal o espacial y la comunicación aleatoria.

La comunicación vecinal se basa en que cada individuo de la población se comunicará con sus vecinos, es decir, recibirá mensajes de los individuos de la población que vivan cerca de él.

La comunicación aleatoria se basa en que los individuos que envían mensajes a otros individuos lo hacen de forma aleatoria, sin importar el lugar que ocupen en la población, ni su vecindad.

Dentro de ambas comunicaciones hay un punto más que deberemos estudiar y es saber quién obtiene el beneficio en la comunicación. Hay dos posibilidades: que en caso de acierto en la comunicación se beneficien de ello tanto el que emite la señal como el que la recibe, llamado beneficio mutuo, o que el beneficio sólo sea para el individuo que ha recibido el mensaje y lo ha entendido, en este caso hablamos de beneficio altruista.

Otro punto importante en el estudio de los sistemas innatos es la reproducción. Es decisivo saber cómo se emparejan unos individuos con otros para dar lugar a nuevas generaciones, y cómo se reparte el material genético que da lugar a los hijos. Para ello estudiaremos cómo influyen en la población distintas técnicas en la reproducción.

En concreto vamos a estudiar dos técnicas: la reproducción aleatoria y la reproducción vecinal o espacial.

La reproducción aleatoria consiste en que el emparejamiento de los individuos de la población de forma aleatoria, es decir, sin atender a su posición en la población, ni relaciones de vecindad.

La reproducción vecinal en cambio basa su funcionamiento en el emparejamiento de individuos vecinos para dar lugar a la siguiente generación de individuos.

Además de todos estos factores de estudio existen una serie de variables que definirán el sistema y que influirán en gran medida en los resultados obtenidos. Dichas variables son: el tamaño de la población que vamos a estudiar, el tamaño total del lenguaje del que cada individuo conocerá una cierta parte, el número de generaciones que queremos que dure la simulación de la evolución, la probabilidad con que queremos que se crucen (reproduzcan) los individuos, y la probabilidad con la que queremos que muten los individuos.

Variando cada uno de estos parámetros podremos ir obteniendo distintos valores y tras el estudio determinaremos cual son los mejores para que la población termine entendiéndose.

### **4.3. DESCRIPCION DEL ALGORITMO EVOLUTIVO PARA INNATOS**

#### INTRODUCCIÓN:

Para la realización del estudio de sistemas innatos de señales hemos elegido un algoritmo evolutivo que se acerque lo máximo posible a la representación real de una población y su lenguaje. Se trata de los Programas Evolutivos, que utilizan representaciones más naturales y más cercanas al dominio del problema.

Antes de empezar tuvimos que decidir la representación del lenguaje que íbamos a utilizar, así como la representación de cada individuo de la población para que fuesen lo más parecidos a lo que entendemos por un lenguaje y por la forma en que cada individuo almacena los conocimientos de que dispone.

Para poder utilizar tamaños de lenguaje muy altos y para evitar errores con el posible orden de concatenación de cadenas de caracteres, decidimos utilizar como lenguaje, el conjunto de los números naturales. Este lenguaje es el que conocerán los individuos y el que ira pasando de unas generaciones a otras.

El tamaño del lenguaje utilizado es uno de los factores a estudio en esta parte del proyecto.

Por otro lado tuvimos que elegir una representación de los individuos que forman la población que se va a estudiar.

Cada uno de ellos está formado por una cadena de genes, que será un subconjunto de los elementos del lenguaje y representará las palabras que conoce dicho individuo de entre todas las palabras existentes.

Además tienen una aptitud, que será donde se vayan guardando el número de aciertos en la comunicación del individuo con el resto de la población y un número de comunicaciones que será el número de intentos de comunicación, tanto los aciertos como los fallos. El proceso de comunicación lo explicaremos en profundidad más adelante.

Cada individuo, además, esta compuesto por una edad que nos indicará el número de generaciones que lleva vivo. Y por último la posición de su padre, que nos será útil en el caso de la reproducción vecinal, para no perder la situación de los individuos dentro de la población de una generación a otra.

## LENGUAJE

El lenguaje que se utiliza en la población es uno de los factores que más influirá en la obtención de unos resultados u otros.

El lenguaje esta compuesto por números enteros consecutivos desde el cero hasta el tamaño del lenguaje, que será una variable de estudio en el sistema.

Cada número entero del lenguaje representa una señal innata de los individuos de nuestra población, sin embargo, se utiliza esta representación para facilitar el manejo de las señales en los procesos de comunicación, reproducción, cruce, etc.

Cada individuo de la población conocerá una parte de este lenguaje y será lo que le sirva para comunicarse con el resto de los individuos de la población, pasando estos conocimientos de padres a hijos.

## POBLACIÓN

Para poder estudiar los distintos aspectos relacionados con la evolución del lenguaje y su ventaja evolutiva, debemos tener una población sobre la cual hacer los experimentos oportunos.

En este caso tenemos una población, compuesta por una serie de individuos, sobre la que simularemos la evolución y cuyo tamaño será una de las variables influyentes en los resultados obtenidos.



Dicha simulación durará un número determinado de generaciones, que al igual que el tamaño de la población, será un factor a estudiar del sistema creado.

## INDIVIDUOS

La población esta formada por un conjunto de individuos, cuya representación será la siguiente:

Los individuos tienen como característica principal su cadena genética, que en este caso será el conjunto de señales del lenguaje que conoce. Al ser un subconjunto del lenguaje total, su representación será también una cadena de números enteros. La longitud de dicha cadena de genes estará directamente relacionada con el tamaño total del lenguaje.

Además cada individuo tiene una aptitud, que será la variable que permita el estudio y diferenciación del comportamiento de los individuos dentro de la población. Dicha aptitud será un contador de los aciertos de comunicación conseguidos por el individuo y será fundamental en la decisión del paso de los individuos a la siguiente generación o su eliminación.

Otra característica de la representación de los individuos es su número de comunicaciones, que será el número de veces que un individuo envía y/o recibe mensajes de otros individuos. Esta variable servirá para el cálculo de la aptitud.

Cada individuo dispondrá también de un atributo edad, que como su propio nombre indica, será el número de generaciones que lleva vivo un individuo.

## ALGORITMO

El algoritmo evolutivo que utilizamos en este caso es un Programa Evolutivo y el esquema general por el que estará compuesto es el siguiente:

Inicio

Generación del lenguaje.

Generación de una población inicial.

Adaptación.

Para i=0 hasta número de generaciones hacer:

Reproducción.

Mutación.

Adaptación.

Evaluación.

Edad ++.

Fin para.

Devolver la población.

Fin.

Ahora definiremos y explicaremos exhaustivamente cada una de las partes de este esquema:

### Generación del Lenguaje:

Como el lenguaje que vamos a tratar esta compuesto por números enteros, simplemente tendremos que ir guardando dichos números desde el cero hasta tamaño\_lenguaje en un vector.

### Generación de la Población Inicial:

Esta parte se encarga de crear la población a partir de la cual empezaremos el estudio y desde la cual se irán modificando los individuos hasta llegar a los resultados del experimento.

Consiste en crear tantos individuos como nos diga la variable tamaño\_población de la siguiente manera:

Lo primero que haremos es inicializar a cero los atributos que definen al individuo, como son aptitud, número de comunicaciones, edad y la cadena de genes, es decir, las palabras que conoce se elegirán aleatoriamente de entre las que contiene el lenguaje total, cerciorándonos de que no se repita ninguna palabra en los genes de un mismo individuo.

La longitud de la cadena de señales que conoce un individuo es fija para todos ellos y se calculará en función al tamaño del lenguaje total.

Todos los individuos generados de esta forma constituirán la población inicial.

### Adaptación:

Esta función se encarga de realizar la comunicación entre los individuos de la población mediante el envío de señales de unos a otros y de comprobar y premiar los intentos de comunicación que llevan a buen puerto, es decir, cuando un individuo envía un mensaje a un segundo individuo y éste entiende esa señal, se premiará el entendimiento.

La comunicación consiste en elegir una palabra al azar de entre las que posee el emisor y comprobar si dicha señal se encuentra entre las que conoce el receptor, es decir, si esa palabra se encuentra entre las que componen su cadena de genes. En caso afirmativo la comunicación ha tenido éxito, en caso contrario la comunicación habrá fracasado.

Analizaremos dos maneras de premiar la comunicación. El primero consistirá en aumentar en uno la aptitud del receptor de la señal cuando la comunicación sea exitosa. Además incrementaremos la variable número de comunicaciones del receptor cada vez que se produzca una comunicación, tanto si es exitosa como si no. Dicho método recibe el nombre de *beneficio altruista*.

La otra forma consiste en que el aumento de la aptitud se produzca tanto en el receptor como en el emisor de la señal, en este caso hablamos de *beneficio mutuo*. La variable número de comunicaciones, en este caso, también se aumentará en uno cada vez que se produzca el envío de la señal, ya sea para llegar a una situación de éxito o de fracaso, tanto en el emisor como en el receptor.

A parte de diferenciar hacia donde va el beneficio en el proceso de la comunicación, también distinguiremos desde qué individuos provienen las señales de comunicación que le llegan a los individuos.

En este sentido podemos distinguir la comunicación vecinal o espacial y la comunicación aleatoria.

La *comunicación vecinal* se basa en que los individuos que envían las señales, para la comunicación con el individuo estudiado en cada momento, serán los vecinos del receptor. En nuestro caso, como la población está representada con un vector de individuos, los emisores serán el vecino de la izquierda y el vecino de la derecha del individuo receptor.

Debido a la representación utilizada los individuos de las fronteras se tratan de forma especial simulando una disposición circular de la población.

En la *comunicación aleatoria*, como su propio nombre indica, los emisores de señales serán individuos de la población tomados al azar, comprobando que el individuo que envía la señal no sea el mismo que la recibe.

Como ocurre en la comunicación vecinal, en cada intento de comunicación se elegirán dos individuos para que envíen mensajes al individuo receptor a estudio.

### Reproducción:

La reproducción es el proceso por el cual, mezclando el material genético de dos individuos (padres), obtenemos otros dos individuos (hijos) nuevos.

La función que se encarga de la reproducción irá recorriendo la población y para cada individuo obtendremos aleatoriamente un valor entre 0 y 1. Si dicho valor es menor que la variable *probabilidad de cruce*, que será una variable a estudio, entonces se realizará el cruce de dicho individuo con el individuo de la población que le corresponda dependiendo del tipo de reproducción utilizada. En caso de que el valor

tomado al azar sea mayor que la probabilidad de cruce, el individuo no se cruzará y se pasará al siguiente.

Vamos a estudiar dos formas de reproducción, dependiendo de los individuos que entren a formar parte de dicha reproducción. Estas formas son: la reproducción vecinal o espacial y la reproducción aleatoria.

La *reproducción vecinal* es en la que los individuos a emparejar son vecinos, es decir, si un individuo se decide cruzar, porque la probabilidad de cruce así lo indica, lo hará con su vecino de la derecha.

El último elemento de la población se trata de manera especial. En caso de ser seleccionado para el cruce, deberá hacerlo con el primero de la población, siempre y cuando éste no haya sido cruzado en esa generación.

La *reproducción aleatoria* es la que elige a los individuos a cruzar de forma aleatoria entre todos los de la población, siempre que no se hayan cruzado ya con otro individuo en esa generación.

Tanto en la reproducción vecinal como en la aleatoria, una vez emparejados los individuos que van a reproducirse, se realiza el *cruce* entre ellos.

Dicho cruce consiste en la creación de dos nuevos individuos desde los dos anteriores. Para ello inicializamos las variables de los individuos hijos, al igual que hacíamos en la creación de la población inicial, es decir, ponemos a cero las variables aptitud, número de comunicaciones y edad de los dos nuevos individuos. Pero en esta ocasión la cadena de genes no se genera aleatoriamente, sino que es una mezcla de las cadenas de los padres.

La distribución de material genético se hace de la siguiente manera:

Primero el padre 1 va repartiendo de forma intermitente sus palabras a los dos hijos, quedando los genes de éstos últimos a la mitad, una vez que termine de repartir el padre.

Padre 1

4	8	10	15	3	1	40	33	23	9
---	---	----	----	---	---	----	----	----	---

Hijo 1

<b>4</b>	<b>10</b>	<b>3</b>	<b>40</b>	<b>23</b>					
----------	-----------	----------	-----------	-----------	--	--	--	--	--

Hijo 2

<b>8</b>	<b>15</b>	<b>1</b>	<b>33</b>	<b>9</b>					
----------	-----------	----------	-----------	----------	--	--	--	--	--

A continuación el padre 2 comienza a repartir sus genes por turnos a sus dos hijos, siempre y cuando no se repita la señal entre las que ya posee el hijo correspondiente. En caso de coincidencia se cambia el turno y se introduce la señal en los genes del hijo que tiene ahora el testigo, y así hasta haber repartido todas las señales de su cadena de genes.

Padre 2

4	7	13	80	33	3	7	5	12	77
---	---	----	----	----	---	---	---	----	----

Hijo 1

4	10	3	40	23	<b>7</b>	<b>80</b>	<b>33</b>	<b>5</b>	<b>77</b>
---	----	---	----	----	----------	-----------	-----------	----------	-----------

Hijo 2

8	15	1	33	9	<b>4</b>	<b>13</b>	<b>3</b>	<b>7</b>	<b>12</b>
---	----	---	----	---	----------	-----------	----------	----------	-----------

### Mutación:

La mutación consiste en el cambio espontáneo del valor de un elemento en la cadena genética de los individuos.

Para ello recorreremos toda la población de individuos y tomamos un valor de forma aleatoria. Si dicho valor es menor que la probabilidad de mutación (variable a estudio), entonces realizaremos un cambio de una de sus señales, siempre y cuando el valor al azar que vayamos a introducir no se encuentre ya entre las señales de que

dispone el individuo en su lenguaje. La posición en la cadena del elemento mutado se elegirá también de forma aleatoria.

Una vez realizada la reproducción y la mutación de la población debemos invocar de nuevo a la función Adaptación, para que haga comunicarse a la nueva población formada de padres e hijos y actualice sus variables aptitud, resultantes de los éxitos o fracasos de la comunicación entre ellos.

### Evaluación:

La función de evaluación es la encargada de seleccionar qué individuos pasarán a la siguiente generación. La estrategia seguida será que pasen sólo los mejores, es decir, los que tengan mejor comportamiento en la comunicación con el resto de la población, o lo que es lo mismo, los que mejor aptitud tengan en ese momento.

Para ello vamos a diferenciar dos funciones de evaluación dependiendo del tipo de reproducción utilizada, es decir, evaluación vecinal o evaluación aleatoria.

Tras la reproducción vecinal realizaremos la correspondiente evaluación, que consistirá en ir eligiendo de entre la población de padres y la de hijos los mejores individuos en cada posición. Esto lo podemos hacer porque en la reproducción vecinal conseguimos mantener la localidad dentro de la población de los hijos respecto a la de los padres.

Por lo tanto la técnica que utilizaremos será recorrer ambas poblaciones y en cada posición seleccionar el mayor de ambas, pasándolo a la población solución, que será la que analicemos en la siguiente generación, o, si es la última iteración, será la población que obtengamos como solución del problema.

En el caso de utilizar una reproducción aleatoria, debemos usar la evaluación aleatoria.

Esta evaluación es más general, y se basa en la mezcla de la población de padres y de hijos y la elección de los X mejores, siendo X el tamaño de la población inicial.

Para la selección de los individuos con mejores aptitudes utilizaremos un algoritmo de ordenación rápida, tras lo cual elegimos la parte de la población resultante correspondiente al tamaño indicado.

Edad:

Antes de pasar a la siguiente generación es necesario incrementar en uno la edad de todos los individuos que se disponen a continuar en la simulación, es decir, los que la función de evaluación ha considerado apropiados para continuar viviendo.

Una vez hecho todo esto nos disponemos a pasar a la siguiente generación, que como indica el esquema repetirá las mismas funciones una y otra vez durante el ciclo de vida de la simulación, para devolver finalmente una población resultado que será de la que obtengamos los resultados del estudios de los sistemas de señales innatas.



#### 4.4. RESULTADOS

Ahora mostraremos los resultados obtenidos en la ejecución de este sistema, estudiando la influencia de ciertas variables en dichos resultados, siendo estas variables las siguientes:

TAM\_POB: Es el tamaño de la población sobre la que haremos la simulación del experimento.

TAM\_LEN: El tamaño del lenguaje total que existe en una población, conociendo los individuos un subconjunto de dicho lenguaje.

NUM\_GEN: Será el número de veces que se repetirá el algoritmo, es decir, el número de generaciones que simula el algoritmo la vida de los individuos de la población.

CRUCE: Es la probabilidad de cruce, es decir, la probabilidad con la que se cruza cada individuo. Esta variable es común para todos los individuos que componen la población.

MUTACIÓN: Es la probabilidad de mutación, es decir, la probabilidad con la que muta algún gen de la cadena que contiene el lenguaje que conoce cada individuo.

ALTRUISTA: La casilla se rellenará con un “SI”, en caso de que el beneficio en la comunicación sea altruista, y con un “NO” en caso de ser un beneficio mutuo.

ALEATORIA: Indica el tipo de comunicación. Si se rellena con un “SI” entonces hablaremos de una comunicación aleatoria y en caso de aparecer un “NO” será que la comunicación es vecinal.

VECINOS: Se está hablando del tipo de reproducción. Si en la casilla correspondiente aparece un “SI” será que la reproducción que estudiemos sea vecinal, en caso de aparecer un “NO” hablamos de una reproducción aleatoria.

Los resultados los obtendremos en las columnas % COMUN, y AP\_MEDIA, que mostrarán respectivamente el porcentaje de comunicación total entre todos los individuos de la población y la aptitud media de la población.

Hemos realizado varias pruebas con los mismos datos de entrada. Cada valor mostrado en la tabla es un promedio de los resultados obtenidos en cada prueba, ya que debido a la aleatoriedad del sistema se consiguen resultados muy variados.

TAM_POB	TAM_LEN	NUM_GEN	CRUCE	MUTAC	ALTRUISTA	ALEATOR	VECINOS	% COINC.	AP MEDIA
10	100	20	0.5	0.01	SI	SI	SI	22.64	49.83
					SI	NO	SI	24.58	31.17
					SI	NO	NO	54.72	81.76
					NO	SI	SI	26.6	39.3
					NO	NO	SI	24.72	43.1
					NO	NO	NO	51.32	57.84
100	100	30	0.5	0.01	SI	SI	SI	25.97	49.08
					SI	NO	SI	24.58	56.31
					SI	NO	NO	45.27	62.57
					NO	SI	SI	27.36	54.61
					NO	NO	SI	25.27	51.54
					NO	NO	NO	62.91	82.39
200	1000	30	0.5	0.1	SI	SI	SI	17.3	33.06

TAM_POB	TAM_LEN	NUM_GEN	CRUCE	MUTAC	ALTRUISTA	ALEATOR	VECINOS	% COINC.	AP MEDIA
200	1000	30	0.5	0.01	SI	NO	SI	17.44	74.86
					SI	NO	NO	20.84	39.95
					NO	SI	SI	17.45	28.68
					NO	NO	SI	17.32	65.71
					NO	NO	NO	21.36	34.45
200	1000	50	0.1	0.05	SI	SI	SI	17.81	35.22
					SI	NO	SI	17.68	78.08
					SI	NO	NO	22.78	42.65
					NO	SI	SI	17.7	28.65
					NO	NO	SI	17.39	68.36
					NO	NO	NO	24.48	38.24
300	1000	30	0.5	0.01	SI	SI	SI	17.09	31.43
					SI	NO	SI	17.06	76.24

TAM_POB	TAM_LEN	NUM_GEN	CRUCE	MUTAC	ALTRUISTA	ALEATOR	VECINOS	% COINC.	AP MEDIA
300	1000	30	0.5	0.01	SI	NO	NO	18.69	38.64
					NO	SI	SI	17.07	29.99
					NO	NO	SI	17.05	62.52
					NO	NO	NO	19.15	30.37
300	1000	50	0.5	0.01	SI	SI	SI	17.26	34.25
					SI	NO	SI	17.15	77.78
					SI	NO	NO	20.1	39.98
					NO	SI	SI	17.28	28.21
					NO	NO	SI	17.06	62.0
					NO	NO	NO	20.78	33.93
300	1000	30	0.6	0.01	SI	SI	SI	17.03	32.08
					SI	NO	SI	17.1	73.08
					SI	NO	NO	18.62	38.24

TAM_POB	TAM_LEN	NUM_GEN	CRUCE	MUTAC	ALTRUISTA	ALEATOR	VECINOS	% COINC.	AP MEDIA
300	1000	30	0.6	0.01	NO	SI	SI	17.23	28.9
					NO	NO	SI	16.96	63.1
					NO	NO	NO	19.29	33.73
300	1000	30	0.5	0.00	SI	SI	SI	17.11	32.83
					SI	NO	SI	17.13	74.3
					SI	NO	NO	18.71	38.26
					NO	SI	SI	17.12	30.01
					NO	NO	SI	17.07	64.6
					NO	NO	NO	17.79	32.54
300	1000	30	0.5	0.05	SI	SI	SI	17.06	34.71
					SI	NO	SI	17.07	75.26
					SI	NO	NO	19.25	40.55
					NO	SI	SI	17.16	29.68

TAM_POB	TAM_LEN	NUM_GEN	CRUCE	MUTAC	ALTRUISTA	ALEATOR	VECINOS	% COINC.	AP MEDIA
300	1000	30	0.5	0.05	NO	NO	SI	16.96	61.97
					NO	NO	NO	19.45	32.2
300	100	200	0.5	0.0	X	X	NO	1.25	100
100	100	200	0.5	0	X	X	NO	99.9	100
100	100	100	0.5	0	NO	X	NO	59	80
100	100	200	0.5	0	NO	NO	SI	20	85

## **4.5. CONCLUSIONES A PARTIR DE LOS RESULTADOS DE INNATOS**

El objetivo de estas pruebas es estudiar en que condiciones converge al lenguaje en una población.

Hemos tomado unos valores para el tamaño de lenguaje comprendidos entre 100 y 1000 y una población de entre 10 y 300 individuos.

Para valores del lenguaje muy altos comparados con el tamaño de la población obtenemos peores resultados ya que hay un mayor rango de palabras de donde elegir, y por tanto mayor probabilidad de fallo en la comunicación.

Existe una diferencia notable entre los resultados obtenidos en una reproducción vecinal con respecto a una aleatoria. Esto es debido a que en la reproducción vecinal, la evaluación se realiza únicamente con el padre e hijo correspondiente, por lo tanto obtenemos en este caso peores resultados.

En las pruebas realizadas con una probabilidad de mutación, aún siendo muy pequeña, hemos obtenido resultados inesperados ya que ésta fomenta demasiado la aleatoriedad de los resultados.

Nos fijamos ahora en el número de generaciones. Podemos observar que para un número de generaciones mayor a 200, una probabilidad de mutación 0 y una comunicación aleatoria (ya que con vecinos como se ha comentado antes los resultados son peores) se obtiene una convergencia total del lenguaje. Esto es algo que se aproxima bastante a la realidad ya que a mayor número de generaciones el lenguaje se irá reduciendo a uno que deberían acabar conociendo todos los individuos de la población.

## **5. EXPERIMENTOS DE APRENDIZAJE DEL LENGUAJE**

### **5.1. ANTECEDENTES APRENDIZAJE**

En el sistema anterior hemos comentado que los seres vivos nacen con la capacidad de enviar y recibir señales. Hemos visto que los sistemas de comunicación innatos sólo se ocupan de la adaptación mediante selección natural.

Pero los conocimientos innatos no son suficientes para establecer la correspondencia entre las señales y los significados, no viene determinada genéticamente. Los seres vivos establecen esta relación mediante un proceso de aprendizaje durante su vida.

Oliphant [5] describe diversos sistemas de aprendizaje mediante los cuales es posible establecer la correspondencia entre señales y significados.

Los métodos de aprendizaje de su estudio que consideramos más interesante son los siguientes:

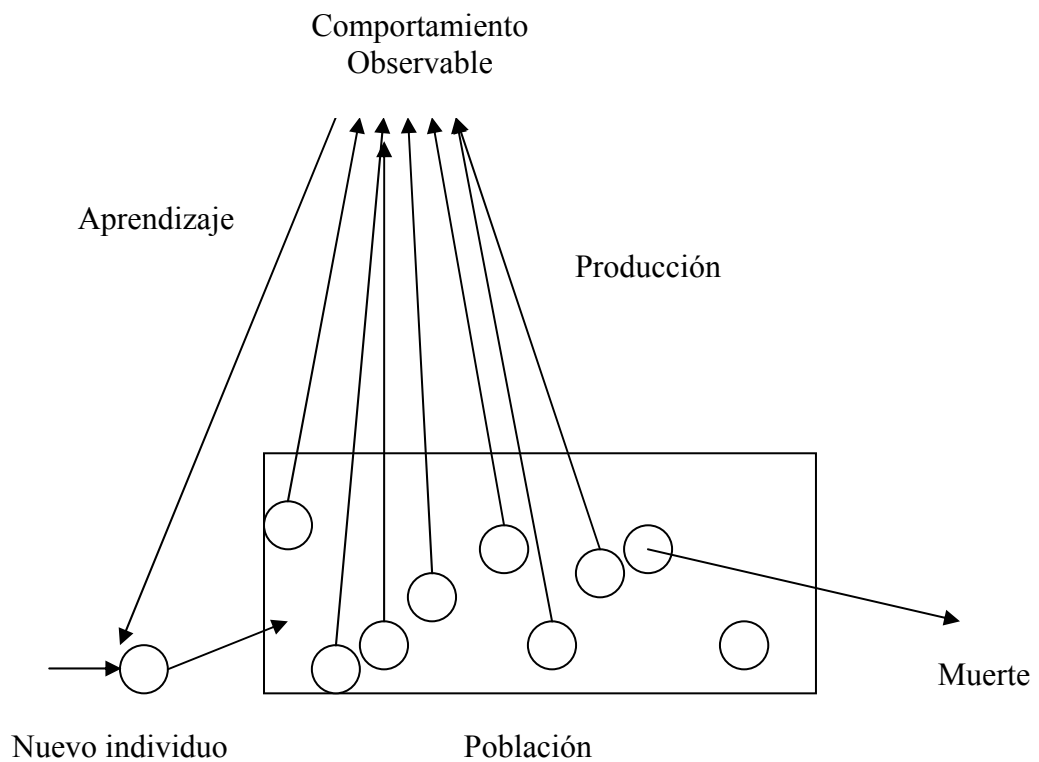
#### **Aprendizaje por reforzamiento**

El sistema de aprendizaje más sencillo es el de prueba-y-error. Usa el mismo método que la selección natural. No es posible construir un sistema de aprendizaje mediante prueba-y-error, ya que en muchos sistemas un fallo en la comunicación puede costar muy caro.



## Aprendizaje por observación

Intuitivamente podemos afirmar que se puede aprender de otros individuos mediante la observación, pero es difícil poder estudiar el efecto del aprendizaje por observación en las poblaciones. Este es el sistema de aprendizaje que utiliza Oliphant en su estudio. El ciclo de aprendizaje de su población queda representado mediante el siguiente diagrama:



Utiliza diversos procedimientos estadísticos de aprendizaje por observación para estudiar con cuál obtiene los mejores resultados. Las técnicas de aprendizaje por observación que estudia son las siguientes:

➤ Aprendizaje por Imitación (Imitate-Choose):

Es el procedimiento de aprendizaje más sencillo, consiste en imitar el comportamiento que se observa en los otros individuos de la población. Oliphant llama a su procedimiento de aprendizaje por imitación Imitate-Choose. El aprendizaje se realiza mandando, para cada significado, la señal que la población suele mandar para ese significado, y recibiendo para esa señal el significado que más frecuentemente recibe la población para esa señal. El procedimiento utilizado para este aprendizaje es el siguiente:

Para cada significado  $\mu$ :

**s.1:** Encontrar la señal  $k_\mu$  para la cual  $S_{\text{smp}}(\mu, k_\mu)$  es máximo.

**s.2:** Asignar  $S_{\text{new}}(\mu, k_\mu)=1.0$ , y asignar  $S_{\text{new}}(\mu, \sigma)=0$  para todo  $\sigma \neq k_\mu$ .

Para cada señal  $\sigma$ :

**r.1:** Encontrar el significado  $\eta_\sigma$  para el cual  $R_{\text{smp}}(\sigma, \eta_\sigma)$  es máximo.

**r.2:** Asignar  $r_{\text{new}}(\eta_\sigma, \sigma)=1.0$ , y asignar  $r_{\text{new}}(\sigma, \mu)=0$  para todo  $\mu \neq \eta_\sigma$ .

Mediante este aprendizaje se refuerza el consenso de la población en la asociación de las señales y los significados. En su estudio Oliphant inicializa la población aleatoriamente y la comunicación es también aleatoria, por lo que obtiene diferentes resultados en distintas ejecuciones. El valor esperado es de  $1/|M|$ .

➤ Aprendizaje Saussurean:

El método de aprendizaje anterior fallaba en su enfoque entre falta de coordinación entre la transmisión y la recepción. Por lo que mediante este aprendizaje se intenta coordinar las funciones de transmisión y recepción. Hurford(1989) llama a la asociación entre señal y significado *señal Saussurean*. Las funciones de envío se construyen igual que el método anterior, pero la de recepción se tendrá que construir forzándose a ser la otra mitad de la *señal Saussurean* de la función de envío.

El procedimiento utilizado es el siguiente:

Para cada significado  $\mu$ :

**s.1:** Encontrar la señal  $k_\mu$  para la cual  $S_{\text{smp}}(\mu, k_\mu)$  es máximo.

**s.2:** Asignar  $S_{\text{new}}(\mu, k_\mu)=1.0$ , y asignar  $S_{\text{new}}(\mu, \sigma)=0$  para todo  $\sigma \neq k_\mu$ .

Para cada señal  $\sigma$ :

**r.1:** Encontrar el significado  $\eta_\sigma$  para el cual  $S_{\text{new}}(\sigma, \eta_\sigma)$  es máximo.

**r.2:** Asignar  $r_{\text{new}}(\eta_\sigma, \sigma)=1.0$ , y asignar  $r_{\text{new}}(\sigma, \mu)=0$  para todo  $\mu \neq \eta_\sigma$ .

Obtiene mejores resultados que con Imitate-Choose, ya que refuerza la coordinación entre los individuos de la población además del consenso. No alcanza los valores óptimos porque no los refuerza indistintamente. Además, al ser la asignación de los valores iniciales a la población aleatoria, los resultados en diferentes ejecuciones son también aleatorios.

➤ Aprendizaje Bayesian:

Con este método se estudia si en vez de imitar el comportamiento de los otros individuos, los aprendices intentan construir un sistema de comunicación basado en la utilidad de la comunicación para la población. En este método los aprendices intentan maximizar su habilidad de comunicación con el mayor número de individuos de la población, es decir, maximizar el acierto en la comunicación.

El procedimiento es el siguiente:

Para cada significado  $\mu$ :

**s.1:** Encontrar la señal  $k_\mu$  para la cual  $R(k_\mu, \mu)$  es máximo.

**s.2:** Asignar  $s_{\text{new}}(\mu, k_\mu)=1.0$ , y asignar  $s_{\text{new}}(\mu, \sigma)=0$  para todo  $\sigma \neq k_\mu$ .

Para cada señal  $\sigma$ :

**r.1:** Encontrar el significado  $\eta_\sigma$  para el cual  $S(\eta_\sigma, \sigma)$  es máximo.

**r.2:** Asignar  $r_{\text{new}}(\eta_\sigma, \sigma)=1.0$ , y asignar  $r_{\text{new}}(\sigma, \mu)=0$  para todo  $\mu \neq \eta_\sigma$ .

Mediante este método Oliphant obtiene los mejores resultados en la comunicación. Demuestra que la población obtenida mediante este método mejorará siempre a la anterior, ya que en cada generación solo se pueden obtener los mismos resultados o mejores, ya que las funciones de envío y recepción maximizan los resultados.

## 5.2. DESCRIPCION APRENDIZAJE

La comunicación entre individuos implica, además de poseer la capacidad de enviar y recibir señales, la habilidad de establecer una correspondencia entre señales y significados. Como hemos comentado anteriormente, esta correspondencia no puede ser adquirida por el individuo a través de su información genética. Será durante su vida, mediante un proceso de aprendizaje, cuando alcance este conocimiento.

Basándonos en estudios anteriores, vamos a desarrollar nuestro propio sistema de aprendizaje. Tomando como referencia el aprendizaje entre los seres vivos, construiremos un sistema donde los individuos sean capaces de adquirir una serie de conocimientos que no son innatos en él.

Kirby [2] expone que la introducción del aprendizaje lleva de una manera natural a la aparición de un concepto interesante: la cultura. Se basa en que la entrada de una agente aprendiz sea la salida de agentes aprendices similares. Es lo que llama *aprendizaje iterativo*. Para facilitar el análisis de los modelos de transmisión cultural propone un marco de aprendizaje iterativo generalizado. Consta de las siguientes componentes:

- Un espacio de significados
- Un espacio de señales
- Uno o más agentes aprendices
- Uno o más agentes adultos

Los agentes adultos son los que tienen el conocimiento del lenguaje, las correspondencias entre las señales y los significados. Los aprendices muestrean pares señal-significado hasta alcanzar un nivel de conocimiento suficiente para ser adulto.

En nuestro sistema hemos utilizado el marco de trabajo propuesto por Kirby para implementar el sistema de aprendizaje.

El lenguaje está compuesto de pares señal-significado. En la población inicial, asignamos a la población adulta un cierto número de pares, los suficientes para que el aprendizaje sea posible. La población inicial de aprendices carece de conocimientos, aunque poseen la capacidad de comunicarse y aprender. En cada iteración preguntan aleatoriamente a los individuos de la población adulta. Cuando el número de pares que conoce un aprendiz es tal que puede ser considerado adulto, este pasa a pertenecer a la población de adultos. Tanto la población de aprendices como la de adultos tienen un número fijo de individuos, que un individuo aprendiz pase a ser adulto implica la eliminación (muerte) de un adulto. Se elimina un adulto al azar, entonces el aprendiz pasa a adulto y se crea un nuevo aprendiz cuyo lenguaje está vacío.

### **5.3. DESCRIPCIÓN DEL ALGORITMO EVOLUTIVO DE APRENDIZAJE**

En este apartado expondremos las características del algoritmo evolutivo de aprendizaje que hemos utilizado para realizar nuestro estudio.

El lenguaje que utilizamos en este algoritmo de aprendizaje está formado por pares señal-significado. La longitud de este lenguaje vendrá determinada como parámetro de entrada de nuestro algoritmo.

Los individuos de nuestro lenguaje guardan como información genética todo el conocimiento aprendido durante su etapa de aprendiz, como pares señal-significado. Esta información no es innata. Los aprendices al nacer no tienen ningún conocimiento, solo poseen la capacidad innata de aprender.

Este experimento trata de observar como son capaces de aprender unos individuos (aprendices) mediante la comunicación (mediante preguntas) con otros individuos adultos. No vamos a estudiar como evoluciona la población utilizando la selección natural, es decir, los individuos de la población no se cruzan. Sólo nos interesa observar las características que hacen que el lenguaje de la población converja.

Para este algoritmo, en vez de tener una única población, tenemos dos: una población de adultos, que son los que poseen el conocimiento del lenguaje, y otra población de aprendices, que son los que deben adquirir el conocimiento. Ambas poblaciones tiene un número fijo de individuos.

Para poder estudiar las características que hacen que el lenguaje que conocen los individuos de la población adulta converja, realizamos sucesivas pruebas variando una serie de datos característicos de la población. Estos datos pueden ser: tamaño de la población, número de generaciones, tamaño del lenguaje, porcentaje de comunicación, esperanza de vida de los individuos y porcentaje de conocimiento para ser adulto.

## **Estructura del Algoritmo de Aprendizaje**

El primer cálculo que se realiza es la inicialización de las variables. Inicializamos el Lenguaje, la población de adultos y la población de aprendices. A continuación ejecutamos el algoritmo de aprendizaje tantas veces como número de generaciones que debe evolucionar la población.

### **Creación del Lenguaje Inicial**

Primero creamos el Lenguaje de la población. Creamos un lenguaje con el tamaño que se indica como entrada del algoritmo y realizamos asignaciones entre señales y significados.

### **Creación de la Población Inicial de Adultos**

La función de inicialización de la población de adultos sería la siguiente:

- Borramos todos los valores guardados anteriormente en la población de adultos.
- Creamos todos los individuos adultos. Para ello realizamos un bucle que se repite tantas veces como número de individuos tenga la población de adultos.

Las operaciones que se realizan son las siguiente:

- Creamos un nuevo adulto, mediante una llamada a una función que inicializa el genoma del adulto.
- Lo mete en la nueva población de adultos.
- Devuelve la población de adultos que hemos creado.

La función que crea un nuevo adulto asigna la información genética de cada adulto. Las operaciones que se realizan en la creación de un nuevo adulto son las que siguen:

- Primero elegimos el número de pares señal-significado que conoce. Es elegido aleatoriamente del tamaño total del lenguaje.



- Para que el conocimiento inicial de los adultos no sea muy pequeño, obligamos a que sea mayor que el 40 por ciento del total. Así podremos contar con una población inicial con el conocimiento suficiente para que los individuos puedan aprender.
- Realizamos un bucle que se repite tantas veces como número de pares que va a conocer el adulto, cuyo valor hemos elegido antes. Dentro del bucle se realizan las siguientes acciones:
  - Se elige un par señal-significado del lenguaje.
  - Este par pasa a formar parte del conocimiento del individuo y se incrementa en una unidad el número de palabras que conoce el adulto.
- Por último se pone a cero la puntuación y la edad del individuo adulto que se está inicializando.
- Se devuelve el adulto que se acaba de crear.

De esta manera quedaría inicializada la población de adultos y se pasaría a crear la población inicial de aprendices.

### **Creación de la Población Inicial de Aprendices**

La inicialización de la población de aprendices se hace de una manera parecida a la inicialización de la población de adultos. El procedimiento realiza los siguientes pasos:

- Borramos todos los valores anteriores que había en la población de aprendices.
- Realizamos el siguiente conjunto de operaciones tantas veces como número de individuos vaya a tener la población:
  - Creamos un aprendiz. Para ello realizamos una llamada a la función que crea los nuevos aprendices.
  - Incluye el aprendiz creado en la población de aprendices.
- Devuelve la población de aprendices que se acaba de crear.

Para la creación del aprendiz tenemos que inicializar el conocimiento de este. Como es un aprendiz no tiene ningún par señal-significado, porque no tiene ningún conocimiento. El esquema que describe la creación de un nuevo aprendiz es el siguiente:

- Borramos toda la información de los genes del individuo. El conocimiento no posee ningún par señal-significado, entonces ponemos el número de pares que conoce a cero.
- También ponemos a cero la edad y la puntuación del aprendiz.
- Una vez creado se devuelve el aprendiz.

Una vez inicializadas la población de Adultos y la población de Aprendices empieza la ejecución del algoritmo de aprendizaje. El algoritmo es un bucle que se repite tantas veces como número de generaciones se introduzca como entrada.

Las instrucciones que se realizan dentro del bucle son las siguientes:

- Primero se realiza la etapa de aprendizaje. Esta se realiza mediante la comunicación entre los aprendices y los adultos.
- A continuación se evalúa el aprendizaje de los individuos.

#### Comunicación entre Aprendices y Adultos

Para cada generación de la población se realiza un turno de aprendizaje de los individuos. Para ello se realizan las siguientes acciones:

- Para cada individuo de la población de aprendices hacemos:
  - Cada aprendiz realiza cierto número de preguntas a los adultos. Este número viene especificado como parámetro de entrada, el porcentaje de comunicación.
  - Se elige una señal aleatoria y si el aprendiz conoce el significado de esa señal se pasa a elegir otra.
  - Si no conoce el significado entonces pregunta a la población de adultos.

- Si algún adulto conoce el significado esta señal entonces el par señal-significado pasa a formar parte del conocimiento del aprendiz.
- Por último aumentamos en una unidad la edad de todos los adultos y los aprendices.

Para saber si un individuo conoce una señal recorremos el genoma del individuo buscando esa señal. El número de pares que conoce viene determinado por la información genética del individuo. Si recorremos todo su genoma y encontramos la señal que le pasamos, entonces devuelve su significado. Si no conoce esa señal devuelve -1.

La función que pregunta a los adultos por una cierta señal que se utiliza en la función de comunicación, se usa para que el aprendiz pregunte a dos adultos aleatorios el significado de la señal que quiere aprender. Si ninguno de ellos sabe el significado devuelve -1. La implementación de la función sigue el siguiente esquema:

- Mientras ningún adulto pueda contestar al aprendiz, éste realiza preguntas a los adultos. Limitamos el número de preguntas a dos adultos.
  - Se elige un adulto al azar entre la población de adultos.
  - Le pregunta al adulto elegido si conoce esa señal.
  - Si conoce el significado de esa señal termina la ejecución de la función y se devuelve el significado.
  - Si no conoce la señal se elige otro adulto.
- Si algún adulto conoce el significado de la señal devuelve ese significado. Si ninguno la conoce devuelve -1.

Una vez se ha terminado la etapa de aprendizaje de esa generación se pasa a evaluar los resultados del aprendizaje.

## Evaluación del Aprendizaje

La función de evaluación comprueba si algún aprendiz supera el conocimiento mínimo para ser adulto. Los pasos que seguimos para comprobar el conocimiento son los siguientes:

- Recorremos toda la población de aprendices para comprobar su conocimiento. Para cada aprendiz realizamos los siguientes pasos:
  - Comprobamos cuantas palabras sabe el aprendiz.
  - Si supera el conocimiento mínimo para ser adulto deja de ser aprendiz.
  - Para ello lo pasamos a la población de adultos. Como el tamaño de las poblaciones es fijo tenemos que eliminar un adulto.
  - Elegimos el adulto de mayor edad y el aprendiz ocupa su lugar.
  - Creamos un nuevo aprendiz y ocupa el lugar del aprendiz que ha pasado a ser adulto.
- Una vez terminada la ejecución se quedan actualizadas la población de adultos y la de aprendices.

Utilizamos una función auxiliar para encontrar el adulto de mayor edad. Para encontrarlo hacemos lo siguiente:

- Guardamos en una variable auxiliar la edad del adulto de mayor edad que hemos encontrado hasta el momento.
- Recorremos la población de adultos. Para cada uno de ellos comprobamos:
  - La edad del adulto.
  - Si su edad es mayor que la que tenemos guardada en la función auxiliar actualizamos el valor de la variable con la del adulto que estamos procesando.
- Devolvemos la posición en la población de adultos del individuo de mayor edad.

Para cada generación se muestran por pantalla los resultados intermedios de la población de adultos y la población de aprendices. Se muestra gráficamente para cada uno de ellos, la edad de los individuos y el número de palabras que conocen.

Una vez se han realizado todas las generaciones de la población se muestran los resultados obtenidos. Para ello calculamos el acierto en la comunicación de los adultos, que será la puntuación de estos. Con ello podemos comprobar si el algoritmo de aprendizaje ha obtenido buenos resultados con los datos introducidos.

### Evaluación de los Adultos

Para comprobar el acierto en la comunicación de la población realizamos los siguientes cálculos:

- Recorremos la población de adultos y para cada adulto realizamos las siguientes acciones:
  - Elegimos tantos adultos de la población como indique el parámetro del porcentaje de comunicación.
  - Tomamos un adulto al azar y se le pregunta si conoce las señales del adulto que se está procesando.
    - Si hay acierto en la comunicación se suma 1 a la puntuación del adulto que se está procesando.
    - Una vez que se ha preguntado por todas las palabras se divide el resultado entre el número de preguntas que se ha hecho.
  - Cuando ya se ha preguntado a los adultos que se han seleccionado, dividimos el resultado entre el número de adultos a los que se ha preguntado. Así obtenemos el valor de la puntuación de ese individuo.
- Una vez terminada la ejecución quedan actualizadas las puntuaciones de cada adulto.

Una vez calculada la puntuación de los adultos se muestra por pantalla. Además de mostrar gráficamente esta información, también mostramos la edad de los adultos y el número de palabras que conoce cada uno.

## 5.4. PRIMER EXPERIMENTO DE APRENDIZAJE

En este apartado vamos a evaluar los resultados obtenidos con la primera versión de nuestro algoritmo de aprendizaje.

En este primer experimento vamos a variar los siguientes datos de entrada del algoritmo para ver los resultados del aprendizaje:

- Tamaño de la población
- Número de generaciones
- Tamaño del lenguaje

Las características concretas de la primera versión del algoritmo son las siguientes:

- En la asignación inicial de conocimiento en la población de adultos hacemos que como mínimo cada adulto conozca un 40 por ciento del total del tamaño del Lenguaje.
- La asignación inicial de pares señal-significado al conocimiento de los adultos no es aleatoria. El lenguaje es distribuido uniformemente entre todos los individuos de la población.
- El número de preguntas que va a hacer cada aprendiz en cada generación es aleatorio. Como mínimo se obliga a que pregunte un cierto número de señales a dos adultos aleatorios de la población.
- En la función de evaluación un aprendiz pasará a ser adulto cuando aprenda más de un 60 por ciento del tamaño del lenguaje.

Realizamos varias ejecuciones del algoritmo anterior tomando como entradas de éste diversos valores significativos y obtuvimos la tabla de resultados de la página siguiente.

Los datos que tomamos son los siguientes: la cantidad media de palabras que conocen los individuos de la población de adultos, la edad media que tienen y el porcentaje de acierto que obtienen en la comunicación con los otros individuos de la población de adultos.



#### **5.4.1. RESULTADOS DEL PRIMER EXPERIMENTO**

Ahora mostraremos los resultados obtenidos en la ejecución de este sistema, estudiando la influencia de ciertas variables en dichos resultados, siendo estas variables las siguientes:

**TAM POBLACIÓN:** Es el tamaño de la población sobre la que haremos la simulación del experimento.

**TAM LENGUAJE:** El tamaño del lenguaje total que existe en una población, conociendo los individuos un subconjunto de dicho lenguaje.

**NUM GENERAC:** Será el número de veces que se repetirá el algoritmo, es decir, el número de generaciones que simula el algoritmo la vida de los individuos de la población.

Los resultados obtenidos son:

**NUM PALABRAS:** Número de palabras que conoce el individuo.

**EDAD:** Edad del individuo.

**% ACIERTOS:** Porcentaje de aciertos en la comunicación del individuo con el resto de la población.

Hemos realizado varias pruebas con los mismos datos de entrada. Cada valor mostrado en la tabla es un promedio de los resultados obtenidos en cada prueba, ya que debido a la aleatoriedad del sistema se consiguen resultados muy variados.

TAM POBLACION	NUM GENERAC.	TAM LENGUAJE	NUM PALABRAS	EDAD	% ACIERTO
50	10	100	Max de 100 palabras Media 75 palabras	Max de 7 años Media de 3	69 %
50	10	200	Max de 200 palabras Media 150 palabras	Max de 6 años Media de 2.6	63 %
50	10	300	Max de 300 palabras Media 230 palabras	Max de 10 años Media de 2.5	65 %
50	20	100	Max de 100 palabras Media 75 palabras	Max de 6 años Media de 4.2	46 %
50	20	200	Max de 200 palabras Media 160 palabras	Max de 7 años Media de 3	53 %
50	20	300	Max de 300 palabras Media 210 palabras	Max de 10 años Media de 2.5	56 %
50	30	100	Max de 100 palabras Media 80 palabras	Max de 8 años Media de 2.6	45 %
50	30	200	Max de 200 palabras Media 145 palabras	Max de 7 años Media de 3	52 %
50	30	300	Max de 300 palabras Media 230 palabras	Max de 6 años Media de 2.4	70 %
100	10	100	Max de 100 palabras Media 75 palabras	Max de 9 años Media de 3.2	47 %
100	10	200	Max de 200 palabras Media 157 palabras	Max de 8 años Media de 2.4	55 %
100	10	300	Max de 300 palabras Media 230 palabras	Max de 10 años Media de 2.1	56 %
100	20	100	Max de 100 palabras Media 76 palabras	Max de 8 años Media de 2.9	65 %
100	20	200	Max de 200 palabras Media 142 palabras	Max de 7 años Media de 3	46 %

TAM POBLACION	NUM GENERAC.	TAM LENGUAJE	NUM PALABRAS	EDAD	% ACIERTO
100	20	300	Max de 300 palabras Media 240 palabras	Max de 13 años Media de 2.3	54 %
100	30	100	Max de 100 palabras Media 73 palabras	Max de 9 años Media de 2.5	49 %
100	30	200	Max de 200 palabras Media 140 palabras	Max de 6 años Media de 2.8	69 %
100	30	300	Max de 300 palabras Media 245 palabras	Max de 10 años Media de 3.1	55 %
150	10	100	Max de 100 palabras Media 75 palabras	Max de 9 años Media de 2.3	73 %
150	10	200	Max de 200 palabras Media 156 palabras	Max de 10 años Media de 2.5	71 %
150	10	300	Max de 300 palabras Media 235 palabras	Max de 9 años Media de 2.2	60 %
150	20	100	Max de 100 palabras Media 72 palabras	Max de 7 años Media de 2	70 %
150	20	200	Max de 200 palabras Media 160 palabras	Max de 9 años Media de 2.4	64 %
150	20	300	Max de 300 palabras Media 230 palabras	Max de 9 años Media de 2	68 %
150	30	100	Max de 100 palabras Media 80 palabras	Max de 7 años Media de 2.3	67 %
150	30	200	Max de 200 palabras Media 150 palabras	Max de 9 años Media de 2.9	58 %
150	30	300	Max de 300 palabras Media 240 palabras	Max de 13 años Media de 3	65 %

#### **5.4.2. CONCLUSIONES DEL PRIMER EXPERIMENTO DE APRENDIZAJE**

Este primer experimento es muy sencillo, ya que tenemos pocos datos de entrada. Mediante las ejecuciones realizadas se pretende observar los valores de entrada del algoritmo bajo los cuales se obtienen los mejores resultados en el aprendizaje.

El algoritmo toma muchas decisiones al azar, por lo que los resultados de las ejecuciones nos dan valores poco predecibles. Consideramos que este azar en el aprendizaje es un fiel reflejo del aprendizaje de los seres vivos ya que es difícilmente predecible como evolucionará un lenguaje en una población.

Observando las tablas vemos que no hay mucha variación entre los resultados obtenidos si mantenemos fijo el tamaño de la población y variamos el tamaño del lenguaje. Aunque suele empeorar el acierto cuanto más grande es el lenguaje. Esto es normal, ya que cuantas más pares tenga el lenguaje de una población mayor es la dificultad en el aprendizaje.

Si miramos ahora, como varía el acierto en la comunicación cuando aumentamos el número de generaciones nos damos cuenta de que cuantas más generaciones tenga el algoritmo peores resultados obtenemos. Esto se debe a que nuestro método de aprendizaje no está diseñado de tal manera que maximice el acierto, ya que las decisiones que tomamos, a la hora de eliminar un adulto, es totalmente aleatoria. No tenemos en cuenta si eliminamos a un adulto con mucho conocimiento, lo elegimos al azar. Esto da, como se ve en nuestras ejecuciones, malos resultados.

Por último nos fijamos en el tamaño de la población. Vemos que los mejores resultados se alcanzan en las poblaciones con tamaños mayores. Cuanto menor sea la diferencia entre el tamaño del lenguaje y la población, incluso si el tamaño de la población es mayor, es más probable que el adulto al que le pregunte el aprendizaje conozca la señal. El aprendizaje es mejor.

Ahora queremos observar si la asignación inicial del lenguaje influye también en el aprendizaje. Realizaremos un segundo experimento con una asignación aleatoria del lenguaje en la población inicial de adultos.

## **5.5. SEGUNDO EXPERIMENTO DE APRENDIZAJE**

En este segundo experimento vamos a comprobar la importancia de la asignación inicial del lenguaje en la población de adultos.

En el experimento anterior la asignación del conocimiento inicial de la población de adultos se realizaba de manera equitativa. De esta forma forzamos a que los pares señal-significado del lenguaje sean conocidos por más de un adulto. Esto lo hacemos para poder asegurar que el conjunto total de pares del lenguaje es conocido por algún adulto, por lo tanto, los aprendices podrán acceder a este conocimiento mediante el proceso de aprendizaje.

Ahora queremos ver qué resultados podemos obtener si la asignación del conocimiento inicial a la población de adultos se realiza de manera aleatoria.

Elegimos el número de pares que conocen los adultos inicialmente de manera aleatoria, pero obligamos a que sea mayor que el 40 por ciento del lenguaje total.

Los pares que conoce cada adulto son elegidos aleatoriamente de entre todos los pares señal-significado que forman el lenguaje de la población.

Introducimos esta variación en el procedimiento de aprendizaje del primer experimento y estudiamos los resultados obtenidos.

Realizamos las mismas ejecuciones que el experimento anterior con los mismos datos y obtenemos la tabla de resultados de la página siguiente:

### **5.5.1. RESULTADOS DEL SEGUNDO EXPERIMENTO**

A continuación mostraremos los resultados obtenidos en la ejecución de este sistema, estudiando la influencia de ciertas variables en dichos resultados, siendo estas variables las siguientes:

**TAM POBLACIÓN:** Es el tamaño de la población sobre la que haremos la simulación del experimento.

**TAM LENGUAJE:** El tamaño del lenguaje total que existe en una población, conociendo los individuos un subconjunto de dicho lenguaje.

**NUM GENERAC:** Será el número de veces que se repetirá el algoritmo, es decir, el número de generaciones que simula el algoritmo la vida de los individuos de la población.

Los resultados obtenidos son:

**NUM PALABRAS:** Número de palabras que conoce el individuo.

**EDAD:** Edad del individuo.

**% ACIERTOS:** Porcentaje de aciertos en la comunicación del individuo con el resto de la población.

Hemos realizado varias pruebas con los mismos datos de entrada. Cada valor mostrado en la tabla es un promedio de los resultados obtenidos en cada prueba, ya que debido a la aleatoriedad del sistema se consiguen resultados muy variados.

TAM POBLACION	NUM GENERAC.	TAM LENGUAJE	NUM PALABRAS	EDAD	% ACIERTO
50	10	100	Max de 100 palabras Media 72 palabras	Max de 7 años Media de 2.3	46 %
50	10	200	Max de 190 palabras Media 130 palabras	Max de 5 años Media de 2.1	45 %
50	10	300	Max de 300 palabras Media 240 palabras	Max de 7 años Media de 2	70 %
50	20	100	Max de 100 palabras Media 74 palabras	Max de 6 años Media de 2.2	44 %
50	20	200	Max de 200 palabras Media 150 palabras	Max de 11 años Media de 1.9	48 %
50	20	300	Max de 300 palabras Media 230 palabras	Max de 5 años Media de 2.1	46 %
50	30	100	Max de 100 palabras Media 79 palabras	Max de 8 años Media de 1.8	72 %
50	30	200	Max de 200 palabras Media 143 palabras	Max de 6 años Media de 2.2	52 %
50	30	300	Max de 300 palabras Media 220 palabras	Max de 7 años Media de 2.3	47 %
100	10	100	Max de 100 palabras Media 72 palabras	Max de 9 años Media de 2.5	71 %
100	10	200	Max de 200 palabras Media 145 palabras	Max de 10 años Media de 2.6	65 %
100	10	300	Max de 300 palabras Media 255 palabras	Max de 7 años Media de 2.5	63 %
100	20	100	Max de 100 palabras Media 69 palabras	Max de 9 años Media de 2.1	68 %
100	20	200	Max de 200 palabras Media 142 palabras	Max de 7 años Media de 2.3	46 %



TAM POBLACION	NUM GENERAC.	TAM LENGUAJE	NUM PALABRAS	EDAD	% ACIERTO
100	20	300	Max de 300 palabras Media 230 palabras	Max de 8 años Media de 2.4	54 %
100	30	100	Max de 100 palabras Media 75 palabras	Max de 9 años Media de 2.2	64 %
100	30	200	Max de 200 palabras Media 150 palabras	Max de 8 años Media de 2.3	45 %
100	30	300	Max de 300 palabras Media 225 palabras	Max de 9 años Media de 2.5	53 %
150	10	100	Max de 100 palabras Media 76 palabras	Max de 8 años Media de 2.9	45 %
150	10	200	Max de 200 palabras Media 155 palabras	Max de 9 años Media de 2.8	71 %
150	10	300	Max de 300 palabras Media 240 palabras	Max de 10 años Media de 2.4	47 %
150	20	100	Max de 100 palabras Media 79 palabras	Max de 8 años Media de 2.6	52 %
150	20	200	Max de 200 palabras Media 158 palabras	Max de 9 años Media de 2.7	47 %
150	20	300	Max de 300 palabras Media 230 palabras	Max de 9 años Media de 2.5	45 %
150	30	100	Max de 100 palabras Media 72 palabras	Max de 9 años Media de 2.7	68 %
150	30	200	Max de 200 palabras Media 155 palabras	Max de 18 años Media de 2.9	72 %
150	30	300	Max de 300 palabras Media 240 palabras	Max de 10 años Media de 2.8	73 %

### **5.5.2. CONCLUSIONES DEL SEGUNDO EXPERIMENTO DE APRENDIZAJE**

En este segundo experimento sólo queremos ver la influencia en el aprendizaje de una asignación aleatoria del lenguaje inicial de la población de adultos.

Para ver la influencia tenemos que compararlo con la tabla del experimento anterior.

La asignación aleatoria del lenguaje hace que los resultados sean más impredecibles, por lo que los datos obtenidos son todavía más aleatorios que en el anterior. Aunque se puede observar, que en líneas generales, las conclusiones que hemos obtenido del experimento anterior también son aplicables en este experimento.

Si comparamos el porcentaje de acierto del experimento anterior con el de éste, vemos que, salvo en contadas excepciones, el aprendizaje empeora con la asignación del lenguaje al azar. Las excepciones se pueden deber a las decisiones al azar que se toman en la ejecución del algoritmo.

El empeoramiento del acierto en la comunicación se puede deber a que en el primer experimento la asignación del lenguaje inicial se hace repartiendo equitativamente el lenguaje entre los adultos, por lo que habrá más probabilidad de que varios adultos conozcan una misma señal. Por lo que el aprendizaje será más fácil.

Con la asignación del lenguaje inicial de la población de adultos del primer experimento vamos a realizar un tercer experimento. Ahora vamos a estudiar la influencia de los siguientes parámetros:

- Porcentaje de Comunicación
- Esperanza de Vida
- Porcentaje de Conocimiento para ser Adulto

## 5.6 TERCER EXPERIMENTO DE APRENDIZAJE

Después de haber estudiado cual es la mejor asignación inicial de la población de adultos, vamos a ver qué factores hacen que mejore el proceso de aprendizaje.

En este experimento vamos a variar los siguientes datos de entrada del algoritmo para ver los resultados del aprendizaje:

- Tamaño de la Población
- Número de Generaciones
- Tamaño del Lenguaje
- Porcentaje de Comunicación
- Esperanza de Vida
- Porcentaje de Conocimiento para ser Adulto

El porcentaje de comunicación indica un tanto por ciento del total del lenguaje que se va a utilizar como medida en el número de comunicaciones que puede realizar cada individuo. Este valor especifica el número de preguntas que puede hacer un aprendiz a la población de adultos en cada generación.

La esperanza de vida media especifica la edad esperada de la muerte de cada individuo adulto. Este valor es utilizado para saber qué individuo va a ser eliminado de la población de adultos cuando un aprendiz alcanza el conocimiento suficiente para ser adulto. Será eliminado entonces el adulto de mayor edad.

Si el valor de la esperanza de vida es cero, se elige el adulto a eliminar al azar. Esta opción se utiliza para comparar la eliminación al azar con la dependiente de la edad.

El porcentaje de conocimiento para ser adulto indica el tanto por ciento del total del tamaño del Lenguaje que tiene que conocer cada aprendiz para pasar a ser adulto.

Mediante este último experimento queremos estudiar cómo evoluciona el aprendizaje de la población variando los datos anteriores.

Como resultado de cada ejecución mostramos el número de palabras que conocen los adultos, la edad que tienen y el acierto en la comunicación. Así podremos observar la eficiencia del aprendizaje y si el lenguaje que conocen los adultos llega a converger en algún momento.

Realizamos varias ejecuciones del algoritmo anterior tomando como entradas del algoritmo diversos valores significativos y obtuvimos la tabla de resultados de la página siguiente.

### 5.6.1. RESULTADOS DEL TERCER EXPERIMENTO

Ahora mostraremos los resultados obtenidos en la ejecución de este sistema, estudiando la influencia de ciertas variables en dichos resultados, siendo estas variables las siguientes:

TamPob: Es el tamaño de la población sobre la que haremos la simulación del experimento.

Nº Generac.: Será el número de veces que se repetirá el algoritmo, es decir, el número de generaciones que simula el algoritmo la vida de los individuos de la población.

TamLenguaj: El tamaño del lenguaje total que existe en una población, conociendo los individuos un subconjunto de dicho lenguaje.

%Comunic.: Porcentaje de la población con la que se comunica cada individuo.

Esp. Vida: Edad máxima que puede sobrevivir un individuo.

%Conocim.: Porcentaje de conocimiento que debe conocer un aprendiz para alcanzar la madurez.

Los resultados obtenidos son:

NumPalabM: Número de palabras que conoce el individuo.

Edad Media: Edad media de la población.

% Acierto: Porcentaje de aciertos en la comunicación del individuo con el resto de la población.

E max Apre.: Edad máxima de los aprendices.

Hemos realizado varias pruebas con los mismos datos de entrada. Cada valor mostrado en la tabla es un promedio de los resultados obtenidos en cada prueba, ya que debido a la aleatoriedad del sistema se consiguen resultados muy variados.

<b>TamPob</b>	<b>N° Generac.</b>	<b>TamLenguaj</b>	<b>%Comunic.</b>	<b>Esp. Vida</b>	<b>%Conocim.</b>	<b>NumPalabM</b>	<b>Edad Media</b>	<b>% Acierto</b>	<b>E max Apre.</b>
150	20	100	25 %	0	30 %	37	6.5	33 %	3
150	20	100	50 %	0	30 %	40	4	37%	1
150	20	100	75 %	0	30 %	43	2	36%	0
150	20	100	25 %	10	30 %	35	4	35%	3
150	20	100	50 %	10	30 %	42	2.3	37%	1
150	20	100	75 %	10	30 %	43	1	36%	0
150	20	100	25 %	20	30 %	36	4	33%	3
150	20	100	50 %	20	30 %	35	2	36%	1
150	20	100	75 %	20	30 %	45	1	35%	0
150	20	100	25 %	0	50%	57	1 de 20 Media: 9	52%	4
150	20	100	50 %	0	50%	65	1 de 16 Media: 5	55%	1
150	20	100	75 %	0	50%	60	1 de 9 Media: 3	47%	1
150	20	100	25 %	10	50%	57	6	52%	4
150	20	100	50 %	10	50%	68	2	55%	1

<b>TamPob</b>	<b>N° Generac.</b>	<b>TamLenguaj</b>	<b>%Comunic.</b>	<b>Esp. Vida</b>	<b>%Conocim.</b>	<b>NumPalabM</b>	<b>Edad Media</b>	<b>% Acierto</b>	<b>E max Apre.</b>
150	20	100	75 %	10	50%	50 (Max 90)	2	47%	1
150	20	100	25 %	20	50%	55(Max 67)	Max: 7 Media: 5	53%	2
150	20	100	50 %	20	50%	65(Max 70)	2	55%	1
150	20	100	75 %	20	50%	55(Max 80)	2	46%	1
150	20	100	25 %	0	75%	80 (Max 100)	Max: 20 Media: 10	71%	6
150	20	100	50 %	0	75%	85 (Max 100)	Max: 18 Media: 8	70%	3
150	20	100	75 %	0	75%	98	Max: 14 Media: 5	80%	1
150	20	100	25 %	10	75%	76	Max:8 Media: 5	70%	6
150	20	100	50 %	10	75%	84	4	69%	2
150	20	100	75 %	10	75%	95	2	75%	1
150	20	100	25 %	20	75%	77	8	74%	7
150	20	100	50 %	20	75%	80	4	68%	2
150	20	100	75 %	20	75%	93	2	79%	1

### **5.6.2. CONCLUSIONES DEL TERCER EXPERIMENTO DE APRENDIZAJE**

Queremos estudiar la influencia que tiene el porcentaje de comunicación, la esperanza de vida de los individuos de la población y el porcentaje de conocimiento que debe tener un aprendiz para ser adulto, en la convergencia de los pares señal-significado de la población adulta final.

Para ello tomamos valores fijos del tamaño del lenguaje (150 individuos), número de generaciones (20 generaciones) y tamaño del lenguaje (100 pares señal-significado).

Hemos tomado un valor de tamaño de la población mayor a la longitud del lenguaje para poder observar mejor las condiciones bajo las cuales los pares de la última generación acaban convergiendo.

Tomamos los resultados que obtenemos con diversos valores y extraemos las siguientes conclusiones:

#### Conclusiones:

Si el valor que damos como esperanza de vida de los individuos es 0, cuando el aprendiz pasa a ser adulto, no tenemos en cuenta la edad de los adultos. Entonces se elimina un adulto aleatorio para introducir ese nuevo adulto en la población.

Esto no tiene una gran influencia en el número de palabras que aprenden los individuos ni en el porcentaje de acierto en la comunicación.

Solo influye en la edad media de la población final de adultos, porque la edad de muerte de un adulto es elegida al azar.

Este resultado es completamente normal, ya que la edad de muerte de un adulto no influye en el aprendizaje, ya que sólo aprenden los individuos que pertenecen a la población de aprendices.



Los datos que más influyen en el proceso de aprendizaje de los individuos son: el porcentaje de comunicación y el porcentaje de conocimiento para ser adulto.

Si el porcentaje de pares señal-significado que debe aprender cada individuo es mayor que el porcentaje de comunicación de los aprendices, entonces estos tardarán más generaciones en ser adultos.

En caso contrario, los aprendices pasarán a ser adultos con menor número de palabras. Encontraremos muy pocos que puedan llegar a conocer la totalidad del lenguaje. Cuanto mayor es la diferencia, se puede observar durante la ejecución, que cada vez los adultos tienen menos conocimiento. Esto hace que cada vez sea más difícil que un adulto pueda dar una respuesta a un aprendiz. Perdemos conocimiento.

Cuanto mayor sea el porcentaje de comunicación mayor es el conocimiento que alcanzan los adultos. Cuantas más preguntas hagan los aprendices tendrán mayor probabilidad de encontrar una respuesta, es decir, de aprender.

Si el porcentaje de conocimiento para ser adulto es muy pequeño, los aprendices pasarán a ser adultos con muy poco conocimiento, lo que hace que el número de pares señal-significado que conocen sea muy pequeño.

Al conocer tan pocos pares el porcentaje de acierto es menor.

Cuanto mayor sea el porcentaje de conocimiento para ser adulto, mejores resultados obtenemos, ya que el aprendiz debe conocer mayor número de palabras para ser adulto.

No se observa que con la variación estos datos, porcentaje de conocimiento, la esperanza de vida y el porcentaje de comunicación, los pares señal-significado que conocen los adultos converjan. Sólo obtenemos un buen porcentaje de aciertos cuando los adultos saben un gran número de pares. En ninguno de los resultados obtenidos vemos que conociendo los adultos pocos pares, de media, el porcentaje de acierto sea grande.

## **6. SISTEMAS DE APRENDIZAJE Y SELECCIÓN NATURAL**

### **6.1. DESCRIPCIÓN DE LOS SISTEMAS DE APRENDIZAJE Y SELECCIÓN NATURAL**

Este sistema que vamos a estudiar reúne las características fundamentales del sistema de señales innatas y del sistema de aprendizaje anteriormente descritos.

El sistema se basa en la creación de una población inicial en la que los individuos tendrán un conocimiento innato, es decir, conocerán un cierto conjunto de señales y sus significados del lenguaje existente, desde el momento en el que nacen.

En este estudio se diferenciarán claramente dos subpoblaciones de la población general, la población compuesta por individuos adultos y la población formada por individuos aprendices.

La población general estará caracterizada por la existencia de familias, en las que se irán agrupando los padres e hijos generados en las distintas generaciones, mientras dure su vida.

La vida de los individuos estará supeditada a la variable esperanza de vida, que será una característica a estudiar por el programa, y que será uno de los atributos que pasen por herencia de padres a hijos.

El objetivo de este sistema es que, durante las generaciones que dure la simulación, los individuos aprendices aprendan de los individuos adultos mediante un sistema de aprendizaje, que consistirá en preguntar directamente a los adultos para que enseñen a los aprendices nuevas palabras (señal - significado) que éstos aún no conocían. Para ello será fundamental la variable del porcentaje de comunicación, que describirá el número de comunicaciones que debe hacer un individuo en cada ciclo de aprendizaje para conseguir palabras nuevas.

Dentro de cada generación se incluirán varios ciclos de aprendizaje de los individuos, para que de una generación a otra les de tiempo a los aprendices a poder aprender un número mínimo de palabras del lenguaje para pasar a ser adultos y poder enseñar ellos a los nuevos aprendices. El número de palabras que han de conocer se

representará como la variable porcentaje de conocimiento, que es otra de las características que heredarán los hijos de los padres.

La forma de generar los aprendices en cada generación será a través de la reproducción que, como ocurría en el sistema de señales innatas, podrá ser vecinal o aleatoria.

Cada vez que se produzca un individuo aprendiz mediante el cruce de dos individuos adultos, el hijo heredará del padre el algoritmo de aprendizaje que deberá aplicar en su proceso de conocimiento de nuevas palabras. Dicho algoritmo estará representado por una serie de características como: la esperanza de vida, el tipo de reproducción, el porcentaje de cruce, el porcentaje de mutación, el porcentaje de conocimiento que necesita un aprendiz para pasar a ser adulto, la familia a la que pertenece, el porcentaje de comunicación y el tipo de comunicación.

Antes de pasar a la siguiente generación se deben trasladar a la población de adultos los individuos aprendices con el conocimiento suficiente como para poder enseñar ellos a los nuevos aprendices de la próxima generación.

Los individuos que pasen de una generación a otra serán los que no hayan superado su esperanza de vida y los que den mejor puntuación en el proceso de comunicación, es decir, los que mejor aptitud obtengan al intentar comunicarse con el resto de la población. Por lo tanto los individuos que mejor se entiendan con el resto de la población.

## **6.2. DESCRIPCIÓN DEL ALGORITMO EVOLUTIVO DE APRENDIZAJE Y SELECCIÓN NATURAL**

### INTRODUCCIÓN:

Para realizar el estudio del sistema de señales innatas con aprendizaje hemos elegido el algoritmo evolutivo que más se aproximaba a la representación real del problema. Por ello, al igual que ocurría en el sistema de señales innatas, hemos elegido un Programa Evolutivo.

Lo primero que tuvimos que elegir fue la representación del lenguaje que iban a reconocer los individuos de la población. En este caso, y al igual que en el sistema de aprendizaje, utilizamos como lenguaje un conjunto de pares. Cada par constará de una señal y su correspondiente significado. La codificación de dichos elementos la haremos con números enteros, para evitar errores en el reconocimiento de las señales, es decir, para evitar errores en el caso de que una señal pudiese tener más de un significado o al contrario, que un significado se pudiese expresar mediante dos señales diferentes.

El tamaño de dicho lenguaje será una de las variables a estudio.

En este sistema, la representación de cada individuo será mayor que la de los sistemas anteriores, ya que se trata de una mezcla de ambas representaciones.

El material genético, es decir, las palabras que conoce, lo guardará como un conjunto de pares señal-significado, al igual que el lenguaje general.

Además tiene una serie de características que identificarán a los individuos y que heredarán de sus padres, como su algoritmo de aprendizaje, y otras que serán definidas directamente como variables de estudio y serán comunes a todos los individuos durante la simulación.

El algoritmo estará basado en el aprendizaje de nuevas palabras por parte de los aprendices, que harán preguntas a los adultos con el fin de poder alcanzar un número mínimo de palabras del lenguaje total y pasar ellos a ser adultos a los que poder preguntar.

Para ello se dispondrá de una serie de iteraciones dentro de cada generación, para que a los aprendices les de tiempo de conocer bastantes palabras.

De una generación a otra sólo pasarán aquellos individuos que mejor se comuniquen con el resto de la población, siempre y cuando no hayan superado su esperanza de vida, ya que los individuos que la rebasen morirán y no pasarán a la siguiente generación aunque tuviesen un conocimiento del lenguaje perfecto.

### LENGUAJE:

El lenguaje que utilizaremos en este sistema será una colección de pares, donde cada par representa una señal y su correspondiente significado.

Las señales y los significados estarán codificados como números enteros, ya que de esta forma no tendremos problemas con la polisemia, es decir, pluralidad de significados de una señal, y la sinonimia, es decir, la pluralidad de señales para un mismo significado.

El tamaño del lenguaje será una de las variables genéricas del sistema desde las cuales estudiaremos los efectos en los resultados.

Cada individuo de la población contendrá un número, no necesariamente igual, de pares significado-señal del lenguaje, para representar su conocimiento.

### POBLACIÓN:

La población de este sistema, y sobre la cual haremos ejecutar la simulación de la evolución, estará dividida en dos grandes subpoblaciones, que son: la población de adultos y la población de aprendices.

Dichas poblaciones serán conjuntos de individuos, pero se tratarán de diferente forma.

Los individuos pertenecientes a la población de aprendices pasarán a la de adultos cuando aprendan lo suficiente. De una generación a otra mantendremos las dos poblaciones, ya que en la población de aprendices pueden quedar individuos que no hayan aprendido lo suficiente, pero que quizás en la siguiente generación lo consigan.

## INDIVIDUOS:

Los individuos que forman las dos poblaciones existentes tienen la misma estructura interna, sin embargo, se tratarán de diferente manera dependiendo de la población en la que se encuentren en ese momento.

El elemento más importante del que disponen los individuos es su cadena de *genes*, que representa el conocimiento de que dispone cada uno. Esta cadena esta compuesta por pares señal-significado, y será un subconjunto del lenguaje total.

También tiene un atributo *aptitud*, que será con lo que podamos medir el buen o mal comportamiento de la comunicación de un individuo con el resto de la población. El *número de comunicación*, como su propio nombre indica, se refiere al número de intentos de comunicación de cada individuo y nos servirá para el cálculo de la aptitud.

Otro atributo importante es la *edad*, que junto con la *esperanza de vida*, nos ayudarán a determinar que individuo ha de morir al acabar una generación. Esto ocurrirá cuando la variable esperanza de vida sea igual que la edad.

En este sistema cada individuo es muy distinto al resto de la población, a diferencia de lo que ocurría en el sistema de señales innatas donde los individuos de una simulación eran prácticamente iguales, salvando su material genético. Por eso disponen de ciertas variables que en el otro sistema tomábamos como genéricas y ahora definen directamente a cada individuo. Estas características son las siguientes: *beneficio*, que nos dice si el beneficio de la comunicación entre dos individuos es mutuo o altruista, *porcentaje de comunicación*, que muestra el número de individuos al que se tiene que preguntar en el proceso de aprendizaje, el *porcentaje de conocimiento*, que será el mínimo número de palabras que deberá saber un individuo aprendiz para poder pasar a la población de adultos, *reproducción*, que podrá ser vecinal o aleatoria como ocurría en el primer sistema que estudiamos, *comunicación*, que nos dirá con quién se comunica cada individuo para calcular su aptitud, podrá ser aleatoria o espacial, la *probabilidad de cruce*, que nos dirá la probabilidad con que se reproducen los individuos, la *probabilidad de mutación*, que será la probabilidad de cambios espontáneos en el material genético de los individuos, y por último la *familia* a la que pertenece, que será una variable muy útil a la hora de la reproducción y la comunicación vecinal o espacial.

### ALGORITMO:

El algoritmo evolutivo que utilizamos en este caso es un Programa Evolutivo y el esquema general por el que estará compuesto es el siguiente:

Inicio

Generación del lenguaje.

Generación de una población inicial.

Para  $i=0$  hasta número de generaciones hacer:

    Comunicación.

    Evaluación.

    Cruce.

    Mutación.

    Para  $j=0$  hasta número de ciclos de aprendizaje hacer:

        Aprendizaje.

    Fin Para.

    Aprendices a adultos.

    Edad ++;

    Matar viejos.

Fin para.

Devolver la población.

Fin.

Ahora definiremos y explicaremos exhaustivamente cada una de las partes de este esquema:

### Generación del Lenguaje:

Como la representación del lenguaje se basa en pares de números enteros, su generación será sencilla. Simplemente debemos generar cada una de las señales como números del cero al tamaño del lenguaje y su correspondiente significado como el valor de la señal más el tamaño del lenguaje. Así nos aseguramos de que no haya ningún significado repetido, ni ninguna señal duplicada.

### Generación de la población inicial:

En este estudio la población inicial es fundamental, ya que sus individuos serán los únicos que tengan material genético innato. Para la realización de dicha población vamos generando tantos individuos como tamaño de población se introduzca en el programa, ya que dicho tamaño es una de las variables a estudio.

Para generar cada individuo lo primero que debemos hacer es inicializar sus atributos aptitud, número de comunicaciones y edad a cero. A cada individuo le asignamos un número de familia consecutivo, que luego se ira heredando en la reproducción, de padres a hijos.

El resto de atributos son variables que se introducen desde el exterior y para evitar que todos los individuos tengan el mismo valor y también la farragosa tarea de ir introduciendo cada variable para cada individuo, decidimos pedir un valor aproximado para cada variable y de manera aleatoria sumar o restar a ese valor otro valor aleatorio.

De esta manera, si por ejemplo para la variable porcentaje de comunicación se introduce un valor 50, entonces de manera aleatoria se decidirá sumar o restar un valor de 0 a 20, siempre y cuando no sobrepasemos los márgenes de acción.

Esto lo haremos para la variables porcentaje\_comunicación, para porcentaje\_conocimiento, para esperanza\_vida, para probabilidad\_cruce y para probabilidad\_mutación.

Los atributos reproducción y comunicación serán, aleatoriamente, vecinales o aleatorios.

El atributo beneficio es general para todos los individuos de la población y se pedirá directamente al exterior del sistema.



Por último nos queda inicializar la cadena de genes de cada individuo de la población con pares señal-significado. Para ello vamos eligiendo valores aleatorios de señales y consultando su significado en el lenguaje. Una vez que tenemos el par, lo añadimos a la cadena genética del individuo. Este proceso se repetirá hasta conseguir un número de pares igual al tamaño inicial de la cadena (lcrom), que será un valor calculado a partir del tamaño del lenguaje. Durante todo este proceso hemos ido comprobando que no se repitan señales en la cadena de genes de un mismo individuo.

Desde esta cadena de genes podemos inicializar la variable número\_palabras de cada individuo con la longitud de dicha cadena, ya que esa variable contendrá el número de señales que conoce un individuo en cada momento de su vida.

### Comunicación:

El proceso de comunicación en este sistema consiste en el envío de señales de unos individuos a otros esperando entenderse. Si hay acierto en la comunicación, es decir, si un individuo envía una señal y el receptor la contiene en su cadena de genes, entonces se premia el éxito. Este premio consiste en el aumento de la aptitud del individuo que recibe la señal y la entiende, en el caso de que la comunicación sea de beneficio altruista, o en el aumento de la aptitud tanto del receptor como del emisor si se trata de una comunicación con beneficio mutuo. Además del aumento de la aptitud, también habrá un aumento de la variable número\_comunicaciones cada vez que se produzca un intento de comunicación, ya sea un éxito o un fallo. El aumento de esta variable será similar al de la aptitud, es decir, aumentamos al receptor si es beneficio altruista o al receptor y al emisor si se trata de comunicación con beneficio mutuo.

La comunicación puede ser de dos tipos: comunicación aleatoria y comunicación vecinal o espacial.

En la comunicación lo que hacemos es que el individuo a estudio en ese momento hará tantas posibles comunicaciones como nos diga la variable porcentaje\_comunicación. Dicha variable es un porcentaje, por lo tanto se referirá a dicho tanto por ciento del tamaño de la población. Meteremos ese resultado en una variable llamada número\_preguntas.

En el caso de la comunicación aleatoria, los individuos con los que se comunica, se eligen aleatoriamente de entre todos los de la población, comprobando que no sea él mismo, porque obviamente no tendría sentido preguntarse una palabra a sí mismo.

En el caso de la comunicación vecinal o espacial, el individuo se comunicará con sus vecinos, elegidos de la siguiente manera:

Con  $X$  vecinos hacia la derecha y con  $X$  vecinos hacia la izquierda, siendo  $X$  la mitad de la variable número\_preguntas.

Al finalizar la comunicación de todos los individuos de la población, se calculará la aptitud de cada uno de ellos en función del número de comunicaciones que ha realizado cada uno.

#### Evaluación:

Tras la comunicación ya tenemos las aptitudes de todos los individuos y ahora debemos ver cuales serán los que pasan a esta generación.

La función de evaluación se encarga de elegir los mejores individuos de acuerdo a su aptitud. Dichos individuos serán sobre los que se trabaje en cada generación de la simulación.

El proceso que realizamos para la elección de los mejores es el siguiente:

Ordenamos la población de acuerdo a su aptitud, mediante un algoritmo de ordenación rápida, por motivos de eficiencia. Una vez que tengamos la población ordenada cogeremos los tamaño\_población mejores.

#### Cruce:

Una vez elegidos los individuos que deben seguir en la simulación, nos disponemos a realizar la reproducción de los individuos adultos para dar lugar a los aprendices.

Para realizar el cruce recorreremos toda la población de individuos adultos. Si la probabilidad\_cruce de cada individuo es mayor que una variable aleatoria, calculada para cada individuo, y dicho individuo no ha sido ya cruzado en esta generación, entonces se realizará el cruce del individuo. El otro individuo involucrado en el cruce se

elegirá de diferente manera si se trata de reproducción aleatoria o si se trata de reproducción vecinal. Sabremos qué tipo de reproducción es observando la variable reproducción del individuo.

Para la reproducción vecinal o espacial el segundo individuo padre se elige aleatoriamente de entre el conjunto de posibles vecinos. Dicho conjunto se crea recorriendo la población y viendo los individuos que tienen como atributo familia uno más o uno menos que el primer padre ya elegido. Estos individuos se incluirán en el conjunto de posibles vecinos siempre y cuando no se hayan cruzado ya en esta generación. Si el tamaño de este conjunto es cero, el primer padre elegido no se cruzará y se pasará al siguiente individuo de la población para realizar el mismo proceso de cruce. En el caso de que haya posibles vecinos con los que cruzarse se elegirá uno aleatoriamente y se realizará la mezcla de las características de cada uno para dar lugar a dos hijos que se introducirán en la población de aprendices.

Los atributos de los nuevos hijos se generarán de la siguiente manera:

Las variables `aptitud`, `número_comunicación`, `edad` y `número_palabras` se inicializarán a cero. El resto de variables se heredarán de los padres. Para cada variable de cada hijo se elegirá al azar de qué padre heredará el valor del atributo. Pudiendo ser así los dos hijos completamente iguales, completamente diferentes o una mezcla de ambos padres. De esta manera hacemos que se parezca lo máximo posible a la realidad.

En el caso de la reproducción aleatoria la elección del segundo padre se hará de forma aleatoria de entre todos los individuos de la población. Dando lugar a dos hijos igual que en la reproducción vecinal y repartiendo el material genético o algoritmo de aprendizaje de la misma manera.

En ambos casos la cadena de pares señal-significado de que disponen los hijos será inicializada al vacío, es decir, no tendrán ningún tipo de conocimiento inicial, a parte de los atributos ya incluidos. Esto lo haremos para comprobar lo que son capaces de aprender durante la simulación del programa.

### Mutación:

El operador de mutación consiste en el cambio espontáneo de algún elemento de la cadena de genes de cada individuo.

Para realizarla recorreremos toda la población y si la variable probabilidad\_mutación de cada individuo es mayor que una variable aleatoria creada también para cada individuo, entonces se muta aleatoriamente un elemento de su cadena de genes. El nuevo elemento que se introduce en lugar del mutado es una señal al azar del lenguaje total y su correspondiente significado, siempre y cuando esa señal no se encuentre ya entre las que conoce el individuo.

### Aprendizaje:

Durante una serie de iteraciones por cada generación se ejecutará la función Aprendizaje, que se realiza X veces, siendo X el número de ciclos que se introduzcan desde el exterior del sistema.

Esta función se realiza varias veces para dar tiempo a los individuos de la población de aprendices de conocer las suficientes palabras para llegar a ser adulto. Aunque no necesariamente en cada generación todos los aprendices pasarán a ser adultos. Algunos tardarán varias generaciones y otros nunca llegarán a serlo, muriendo antes porque su edad supere su esperanza de vida.

El número de preguntas que hará cada aprendiz en cada ciclo de aprendizaje se calculará en función del porcentaje de comunicación asociado al individuo y del tamaño total de la población de adultos.

Por tanto, para realizar el aprendizaje, recorreremos los individuos de la población de aprendices. Y cada vez que tenga que preguntar generaremos una señal aleatoria de las existentes en el lenguaje. Si el individuo no conoce ya esa señal entonces se dispondrá a preguntársela a dos individuos de la población de adultos seleccionados aleatoriamente de entre toda esa población. Si alguno de ellos conoce la señal y su significado, se lo comunicará al aprendiz y éste la incluirá entre sus palabras conocidas.

### Aprendices a adultos:

Una vez terminados los ciclos de aprendizaje de cada generación, debemos hacer que los aprendices, que han aprendido un número de palabras mayor que el número que indica su variable *porcentaje de conocimiento* sobre el lenguaje total, pasen de la población de aprendices a la población de adultos. Una vez que pasan a la población de adultos serán borrados de la de aprendices y en la siguiente generación podrán ser ellos los que enseñen a los nuevos aprendices y a los que se quedaron en la población de aprendices porque no conocían las suficientes palabras.

### Edad:

Antes de pasar a la siguiente generación debemos incrementar en uno la edad de los individuos de la población de adultos y los de la población de aprendices.

### Matar viejos:

Una vez incrementada la edad de todos los individuos deberemos simular la muerte de los que hayan alcanzado una edad igual que su variable *esperanza de vida*. Para ello iremos recorriendo ambas poblaciones y quitando los individuos en los que ocurra esto.

Una vez realizados todos estos pasos continuaremos con la siguiente generación. El número de generaciones que hay que repetirlo será una variable introducida externamente y que será uno de los factores de estudio del sistema.

Una vez terminada la iteración de generaciones devolveremos la población resultante desde la que obtendremos los resultados del estudio de la simulación del sistema de señales innatas con aprendizaje.

### 6.3. RESULTADOS

Ahora mostraremos los resultados obtenidos en la ejecución de este sistema, estudiando la influencia de ciertas variables en dichos resultados, siendo estas variables las siguientes:

TamPob: Es el tamaño de la población sobre la que haremos la simulación del experimento.

Nº Generac.: Será el número de veces que se repetirá el algoritmo, es decir, el número de generaciones que simula el algoritmo la vida de los individuos de la población.

NºCiclos: Número de ciclos de aprendizaje dentro de cada generación.

Tam\_Leng: El tamaño del lenguaje total que existe en una población, conociendo los individuos un subconjunto de dicho lenguaje.

Esp. Vida: Edad máxima que puede sobrevivir un individuo.

%Comun.: Porcentaje de la población con la que se comunica cada individuo.

%Conoc.: Porcentaje de conocimiento que debe conocer un aprendiz para alcanzar la madurez.

Prob\_cruce: Es la probabilidad de cruce, es decir, la probabilidad con la que se cruza cada individuo. Esta variable es común para todos los individuos que componen la población.

Prob\_Mutac: Es la probabilidad de mutación, es decir, la probabilidad con la que muta algún gen de la cadena que contiene el lenguaje que conoce cada individuo.

Beneficio: Tipo de beneficio en la comunicación. Puede ser mutuo o altruista.

Los resultados obtenidos son:

%ComTotal: Porcentaje de aciertos en la comunicación entre todos los individuos.

AptMedia: Es la media de la aptitud de todos los individuos de la población.

Hemos realizado varias pruebas con los mismos datos de entrada. Cada valor mostrado en la tabla es un promedio de los resultados obtenidos en cada prueba, ya que debido a la aleatoriedad del sistema se consiguen resultados muy variados.

TamPob	NºGenerac.	NºCiclos	Tam_Leng	Esp. Vida	%Comun.	%Conoc.	Prob_cruce	Prob_Mutac	Beneficio	%ComTotal	AptMedia
10	20	5	100	10	50	60	0,5	0,01	Mutuo	14.95	0.222
10	20	5	100	10	50	60	0,5	0,01	Altruista	21.97	0.212
10	20	5	10	10	50	60	0,5	0,01	Mutuo	0.0	0.0
10	20	5	10	10	50	60	0,5	0,01	Altruista	8.3	0.05
10	20	5	100	10	50	60	0,5	0,0	Mutuo	18.93	0.16
10	20	5	100	10	50	60	0,5	0,0	Altruista	43.07	0.14
10	20	5	10	10	50	60	0,5	0,0	Mutuo	0.0	0.0
10	20	5	10	10	50	60	0,5	0,0	Altruista	0.0	0.0
10	2	5	50	10	50	60	0.5	0.0	Mutuo	15	0.08
10	2	5	50	10	50	60	0.5	0.0	Altruista	10	0.09
10	2	5	50	10	50	60	0.5	0.01	Mutuo	20.58	0.13
10	2	5	50	10	50	60	0.5	0.01	Altruista	12.5	0.14
10	20	5	50	10	50	60	0.5	0.01	Mutuo	34.01	0.34
10	20	5	50	10	50	60	0.5	0.01	Altruista	21.67	0.35

<b>TamPob</b>	<b>N°Generac.</b>	<b>N°Ciclos</b>	<b>Tam_Leng</b>	<b>Esp. Vida</b>	<b>%Comun.</b>	<b>%Conoc.</b>	<b>Prob_cruce</b>	<b>Prob_Mutac</b>	<b>Beneficio</b>	<b>%ComTotal</b>	<b>AptMedia</b>
10	20	5	50	10	50	60	0,5	0,01	Mutuo	30.25	0.12
10	20	5	50	10	50	60	0,5	0,01	Altruista	33.51	0.25
10	5	5	100	10	50	60	0,5	0,01	Mutuo	13.12	0.11
10	5	5	100	10	50	60	0,5	0,01	Altruista	11.36	0.14
50	10	5	100	10	50	60	0.5	0.01	Mutuo	64	0.38
100	5	5	100	10	50	60	0.5	0.1	Mutuo	22.12	0.21
100	5	5	100	10	50	60	0.5	0.1	Altruista	28.68	0.24
50	40	10	100	15	50	60	0.5	0.0	Mutuo	89.7	0.43
50	40	10	100	15	80	80	0.5	0.0	Mutuo	86.8	0.48
100	15	5	100	5	50	60	0.8	0.0	Mutuo	83	0.46
50	10	10	100	10	50	60	0.8	0.0	Mutuo	87	0.50
50	10	10	100	10	50	60	0.8	0.0	Altruista	83.8	0.44
50	10	20	100	10	50	60	0.8	0.0	Altruista	91.33	0.57
50	10	30	100	10	50	60	0.8	0.0	Altruista	90.87	0.51



<b>TamPob</b>	<b>NºGenerac.</b>	<b>NºCiclos</b>	<b>Tam_Leng</b>	<b>Esp. Vida</b>	<b>%Comun.</b>	<b>%Conoc.</b>	<b>Prob_cruce</b>	<b>Prob_Mutac</b>	<b>Beneficio</b>	<b>%ComTotal</b>	<b>AptMedia</b>
50	10	30	100	10	50	60	0.8	0.0	Mutuo	91.46	0.46
50	20	30	100	5	50	60	0.8	0.0	Mutuo	92.49	0.44
100	10	5	100	10	50	60	0.5	0.0	Mutuo	73.9	0.63
100	10	5	100	10	50	60	0.5	0.0	Altruista	82.3	0.49
100	10	5	100	10	50	80	0.5	0.0	Altruista	0.0	0.0
100	10	5	100	10	50	80	0.5	0.0	Mutuo	0.0	0.0
100	10	5	100	10	80	80	0.5	0.0	Mutuo	0.0	0.0
100	10	10	100	10	80	80	0.5	0	Mutuo	78	0.44
100	20	10	100	10	80	80	0.5	0.0	Altruista	76	0.39
100	20	10	100	5	80	30	0.5	0.0	Mutuo	78	0.42
10	20	5	100	10	50	60	0.5	0.0	Mutuo	26.6	0.1
10	20	5	100	10	50	60	0.5	0.0	Altruista	33	0.13
50	10	3	100	10	50	60	0.5	0.01	Mutuo	0.0	0.0
50	10	3	100	10	50	60	0.5	0.01	Altruista	0.0	0.0

## 6.4. CONCLUSIONES APRENDIZAJE Y SELECCIÓN NATURAL

Una vez obtenidos los resultados del experimento mostraremos las conclusiones a las que hemos llegado con su análisis.

Los resultados que obtendremos para cada variación de los factores de estudio serán el *porcentaje de coincidencias total*, que será el acierto de la comunicación de cada individuo con el resto de la población, y la *aptitud media* de la población.

El algoritmo toma muchas decisiones al azar, por lo que los resultados de las ejecuciones nos dan valores poco predecibles. Consideramos que este azar en el aprendizaje es un fiel reflejo del aprendizaje de los seres vivos ya que es difícilmente predecible como evolucionará un lenguaje en una población.

Una de las conclusiones a las que llegamos estudiando las tablas es que si tenemos mucho porcentaje de conocimiento necesario para que un aprendiz pase a ser adulto y también mucho porcentaje de comunicación, es decir, cada individuo realiza muchas consultas a los adultos para aprender, pero existen pocos ciclos de aprendizaje en cada generación, la población tiende a desaparecer.

Una población desaparece cuando sus individuos no aprenden lo suficiente para pasar a la población de adultos y mueren debido a que su edad llega a ser igual que su esperanza de vida.

En el caso anterior pero con mayor número de ciclos de aprendizaje por generación obtenemos resultados mucho mejores.

En general cuando los ciclos de aprendizaje son muy pocos se tiende a perder la población.

En el caso de tener mucho más porcentaje de conocimiento que de comunicación observamos que los resultados empeoran, es decir, el lenguaje final converge menos.

Respecto a la mutación, tenemos que aunque sea muy pequeña hace que se varíen mucho los genes de los individuos, y se obtienen valores inesperados. Por tanto la mutación crea demasiada aleatoriedad.

Si estudiamos el lenguaje veremos que para tamaños muy pequeños la población desaparece. Según vamos aumentando el tamaño los resultados que vamos obteniendo van siendo mejores.

A la hora de estudiar el tamaño de la población veremos que ocurre lo mismo que con el tamaño del lenguaje, que cuanto mayor sea mejores resultados se obtienen.

Cuando tenemos muy poco porcentaje de comunicación, los individuos irán desapareciendo porque no serán capaces de aprender.

## **7. IMPLEMENTACIÓN Y MANUAL DE USUARIO**

### **7.1. IMPLEMENTACIÓN SISTEMA INNATO DE SEÑALES**

Para la realización de la implementación de este sistema hemos utilizado la herramienta C++ Builder.

Las estructuras de datos utilizadas han sido las siguientes:

#### **Estructura de datos para el Lenguaje:**

Para el tipo lenguaje utilizamos un vector de enteros, en el que cada uno simulará una señal de las existentes entre las señales innatas:

```
typedef vector<int> TLenguaje;
```

#### **Estructura de datos para los Individuos:**

Para el tipo individuo utilizamos un registro que constará de los siguientes campos:

El campo *genes* que será del tipo TLenguaje anteriormente definido, el campo *aptitud* que será un número real, y *num\_comunicacion*, *posPadre* y *edad* que son campos con valores enteros:

```

typedef struct
{
    TLenguaje genes;

    double aptitud;

    int num_comunicacion;

    int posPadre;

    int edad;

} TIndividuo;

```

### **Estructura de datos de la Población:**

Para definir el tipo de la población utilizaremos un vector de individuos:

```

typedef vector<TIndividuo> TPoblacion;

```

Las clases que forman parte de este sistema son las siguientes:

- Pantalla principal: Pide los datos de entrada necesarios para la ejecución del algoritmo y se los pasa al programa evolutivo. Cuando finaliza la ejecución muestra los resultados que hemos obtenido.
- Clase Principal: En esta clase tenemos la implementación del programa evolutivo, además del Lenguaje, la población de adultos y la población de aprendices.
- Clase Individuo: Contiene la estructura de datos de los individuos de la población.

## 7.2. MANUAL DE USUARIO DE SISTEMA INNATO DE SEÑALES

La pantalla que aparece al ejecutar nuestro programa es la siguiente:

**Sistemas Innatos de Señales**

**DATOS DE LA POBLACIÓN**

Tamaño de la Población: 50      Número de Generaciones: 20

Probabilidad de Cruce: 0.5

Probabilidad de Mutación: 0.01      Tamaño Lenguaje: 100

**BENEFICIO COMUNICACIÓN:**

☒ Beneficio Altruista

☐ Beneficio Mutuo

**SELECCIÓN COMUNICACIÓN:**

☒ Por Parejas Aleatorias

☐ Por Selección Espacial

**REPRODUCCIÓN:**

☐ Con vecinos

☒ Aleatoriamente

**Aceptar Datos    Nuevo Cálculo    Salir de la Aplicación**

**SOLUCIÓN:**

Porcentaje de coincidencias total:

Aptitud Media de la Población Final:

**Aptitud    Num comunicaciones    Edad**

En esta pantalla se introducen los datos que especifican los parámetros de entrada del algoritmo evolutivo del sistema innato de señales. Para poder ejecutar el programa tenemos que rellenar todas las casillas de los datos de la población. Si alguna de ellas se queda sin rellenar, se mostrará un mensaje de error.

La probabilidad de cruce y la probabilidad de mutación se rellenarán con un valor comprendido entre 0 y 1. Para los valores decimales utilizaremos una “,”.

Si no marcamos ningún valor de beneficio de comunicación, ni de selección en la comunicación, ni de reproducción, los valores que se toman por defecto son beneficio altruista, comunicación por parejas aleatorias y reproducción aleatoria.

Para comenzar la ejecución del algoritmo evolutivo, una vez se han introducido los datos, se pulsa el botón **Aceptar Datos**.

Cuando termine la ejecución del algoritmo entonces se mostrará la siguiente pantalla:

The screenshot shows a software window titled "Sistemas Innatos de Señales". It contains several sections for configuring and displaying results of an algorithm.

**DATOS DE LA POBLACIÓN**

Tamaño de la Población:	50	Número de Generaciones:	20
Probabilidad de Cruce:	0,5		
Probabilidad de Mutación:	0,01	Tamaño Lenguaje	100

**BENEFICIO COMUNICACIÓN**

☒ Beneficio Altruista  
☐ Beneficio Mutuo

**SELECCIÓN COMUNICACIÓN**

☒ Por Parejas Aleatorias  
☐ Por Selección Espacial

**REPRODUCCIÓN**

☐ Con vecinos  
☒ Aleatoriamente

**Buttons:** Aceptar Datos, Nuevo Cálculo, Salir de la Aplicación

**SOLUCIÓN:**

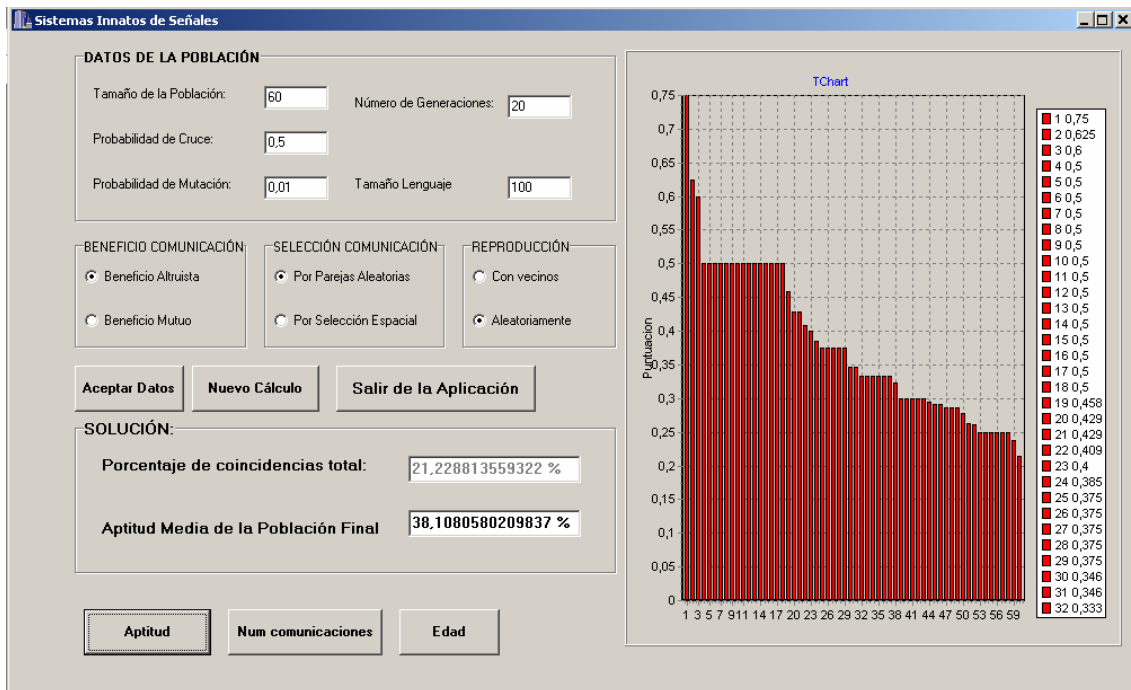
Porcentaje de coincidencias total:	22.7704081632653 %
Aptitud Media de la Población Final	46.5941852007641 %

**Buttons:** Aptitud, Num comunicaciones, Edad

En ella se muestran los valores de porcentaje de coincidencias total en la comunicación y la aptitud media de la población final.

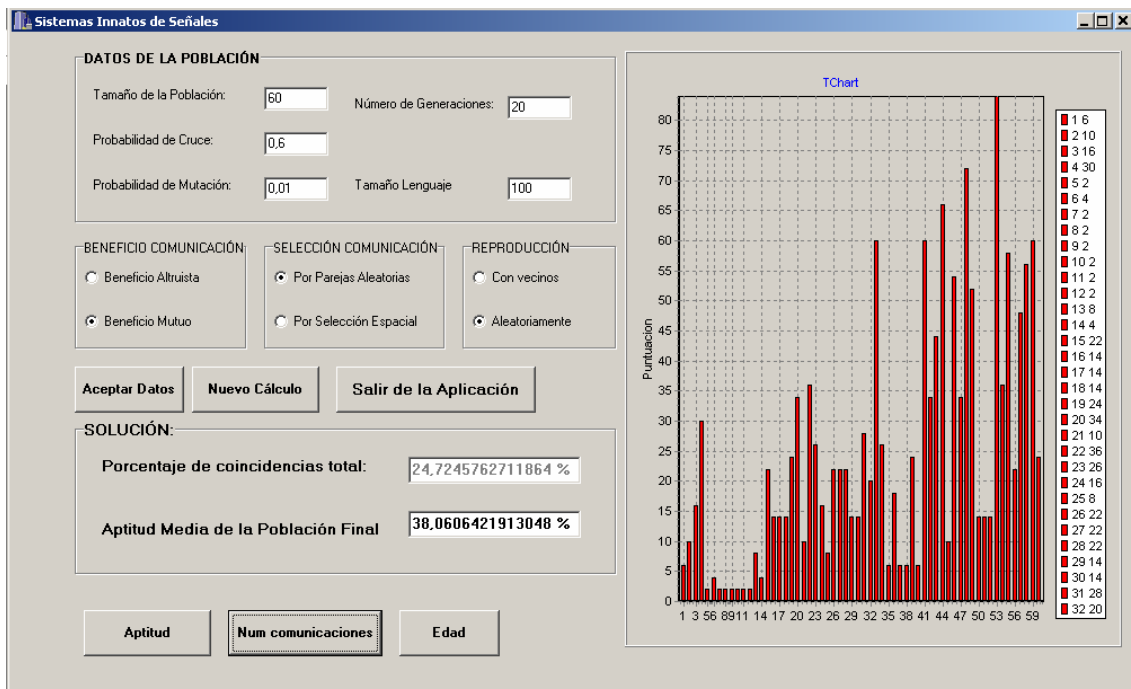
Si queremos ver con más detalle los valores de la ejecución podemos pulsar una serie de botones.

Si pulsamos el botón de **Aptitud** nos aparecerá la siguiente gráfica:



En ella se muestra el valor de la aptitud de cada individuo. Así podremos observar, ordenadas de mayor a menor, las aptitudes de cada individuo de la población.

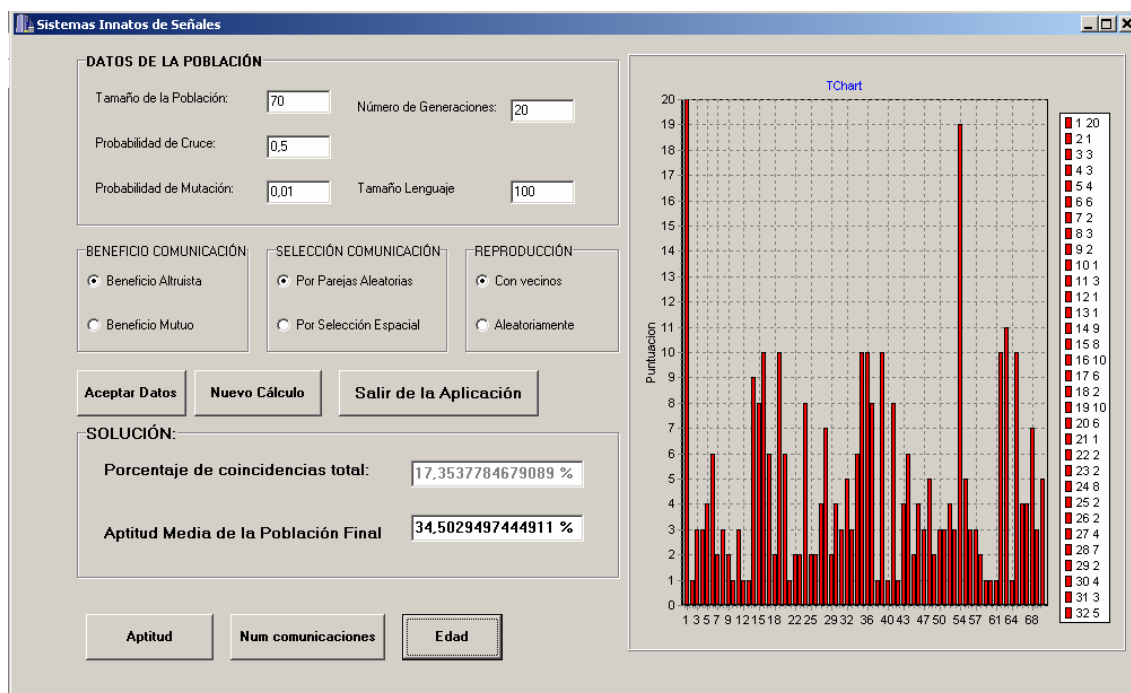
Si pulsamos el botón **Num comunicaciones** nos mostrará la siguiente gráfica:



En la gráfica se muestra, para cada individuo el número veces que se ha comunicado.



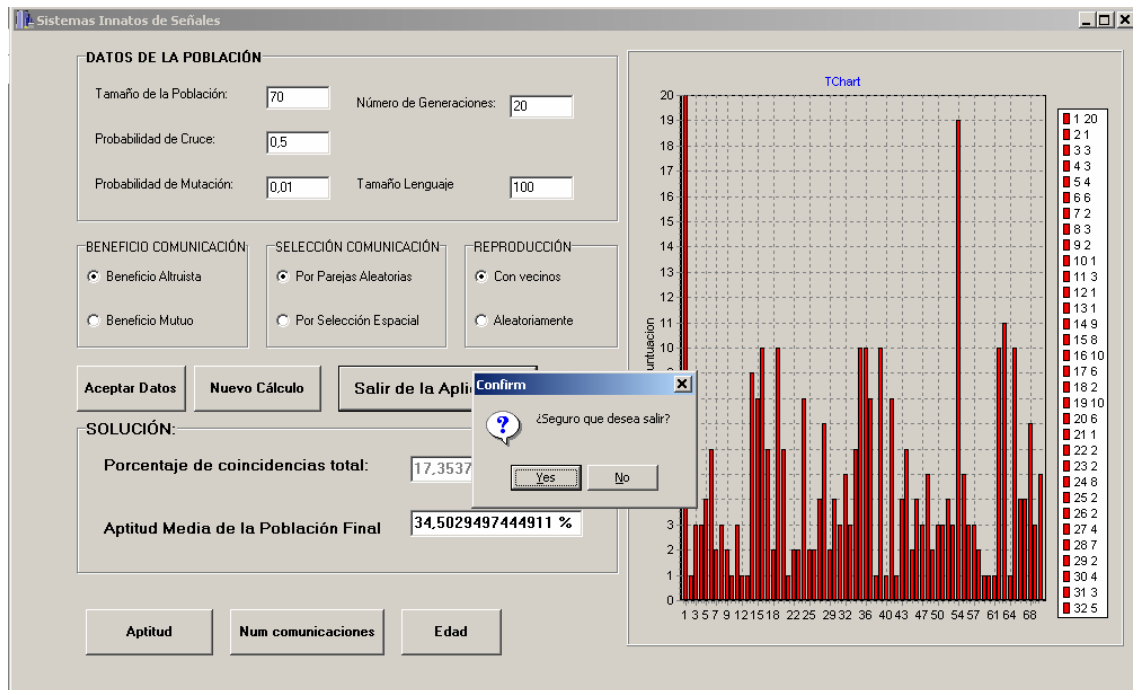
Si pulsamos el botón **Edad** nos mostrará lo siguiente:



Así podremos observar la edad de cada individuo, para poder apreciar aquellos que tienen mayor éxito en la comunicación, ya que han sobrevivido a más generaciones.

Si queremos realizar un nuevo cálculo pulsaremos el botón **Nuevo Cálculo**, entonces se borrarán todos los campos de los datos de entrada de la anterior ejecución y también las gráficas.

Para poder salir de la aplicación se debe pulsar el botón **Salir de la Aplicación**. Se mostrará un mensaje de confirmación. Si se acepta se cerrará la aplicación.



### 7.3. IMPLEMENTACIÓN EXPERIMENTO DE APRENDIZAJE DEL LENGUAJE

La implementación del algoritmo de aprendizaje la hemos realizado con la herramienta C++ Builder.

Primero vamos de describir en detalle la estructura de datos del Lenguaje, de los individuos y de la población.

#### Estructura de datos del Lenguaje

El Lenguaje es el conjunto de todo el conocimiento que puede llegar a conocer la población de individuos.

Definimos la estructura TPares:

```
typedef struct  
  
    {  
  
        int senal;  
  
        int significado;  
  
    }TPares;
```

Esta estructura guarda la relación entre señal y significado.

El Lenguaje es una lista de TPares:

```
typedef vector<TPares> TLenguaje;
```

En esta estructura guardamos los pares señal-significados del Lenguaje de toda la población o en la estructura genética del individuo para guardar el lenguaje que conoce.

### **Estructura de datos de Individuo**

Esta estructura describe la información genética de cada individuo, donde se guardará el lenguaje que conoce, cuantos pares conoce, su edad y su puntuación.

La estructura de individuo es la siguiente:

```
typedef struct  
  
    {  
  
        TLenguaje genes;  
  
        int edad;  
  
        double puntuacion;  
  
        int numPalabras;  
  
    } TIndividuo;
```

El conocimiento de cada individuo solo se verá incrementado en su fase de aprendizaje, una vez haya alcanzado su madurez, el lenguaje que conoce no variará.

### **Estructura de datos de la Población**

En esta estructura de datos guardamos una lista de individuos. La declaración sería la siguiente:

```
typedef vector<TIndividuo> TPoblacion;
```

En nuestro experimento la población está dividida en dos subpoblaciones, la de adultos, que son los individuos que tienen el conocimiento, y la de aprendices.

Las clases que forman parte de nuestro sistema son las siguientes:

- Pantalla principal: Pide los datos de entrada necesarios para la ejecución del algoritmo y se los pasa al algoritmo de aprendizaje. Cuando finaliza la ejecución muestra los resultados que hemos obtenido.
- Pantalla Intermedia: Muestra los resultados intermedios del aprendizaje. Para cada generación muestra los datos de la población de aprendices y la de adultos.
- Clase Principal: En esta clase tenemos la implementación del algoritmo evolutivo, además del Lenguaje, la población de adultos y la población de aprendices.
- Clase Individuo: Contiene la estructura de datos de los individuos de la población.

## 7.4. MANUAL DE USUARIO DE EXPERIMENTO DE APRENDIZAJE DEL LENGUAJE

Cuando ejecutamos la aplicación nos aparece la siguiente pantalla:

**Aprendizaje del Lenguaje**  
Nuevo Cálculo Salir de la Aplicación

**DATOS DE LA POBLACIÓN:**

Tamaño de la población: 100

Número de generaciones: 20

Tamaño Lenguaje: 200

% Comunicación : 50

Esperanza de Vida : 10

% Conocimiento para ser Adulto : 75

Aceptar Datos

**Número de Palabras que Conoce**

Número de Palabras

0

0

Individuos

**Edad de la Población**

Edad

0

0

Individuos

**Porcentaje Acierto Comunicación**

Porcentaje Acierto

0

0

Individuos

En ella tendremos que introducir los valores de entrada de la población que vamos a calcular. Si alguno de los campos de los datos de la población se queda en blanco, no se podrá realizar la ejecución y se mostrará un mensaje de error.

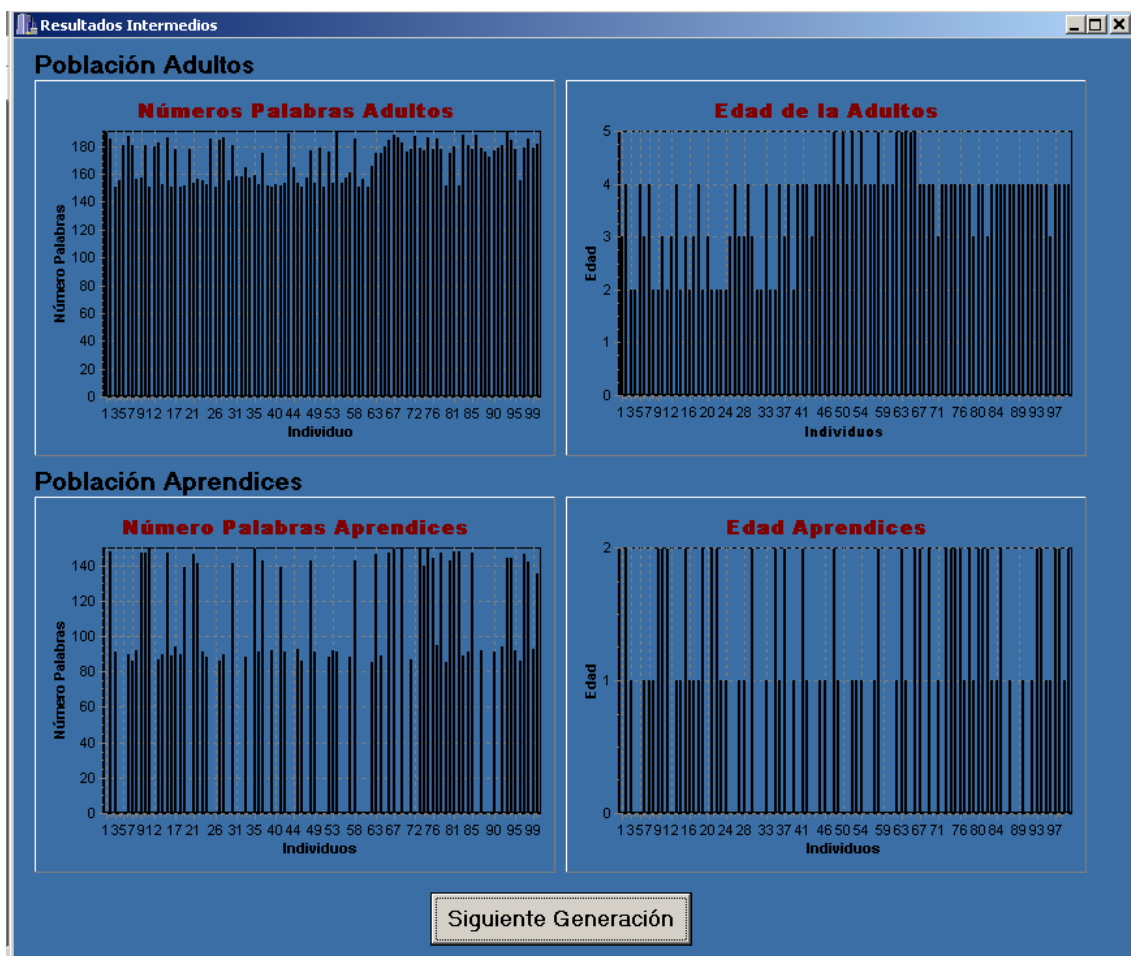
Si los datos de los campos de porcentaje de comunicación, esperanza de vida y porcentaje de conocimiento para ser adulto se quedan a cero estos tomarán los siguientes valores por defecto:

- Porcentaje de Comunicación: será elegido aleatoriamente, pero siempre será superior al 50 por ciento del lenguaje.
- Esperanza de Vida: el adulto que muere al introducir nuevo adulto a la población será elegido al azar.

- Porcentaje de Conocimiento para ser adulto: el aprendiz tendrá que saber más de un 60 por ciento del lenguaje para ser considerado adulto.

Para iniciar la ejecución del programa se pulsará el botón **Aceptar Datos**.

Una vez pulsado el botón se iniciará la ejecución del algoritmo de aprendizaje. Para poder observar mejor como avanza el aprendizaje de la población de aprendices, se muestra en cada generación la siguiente pantalla:



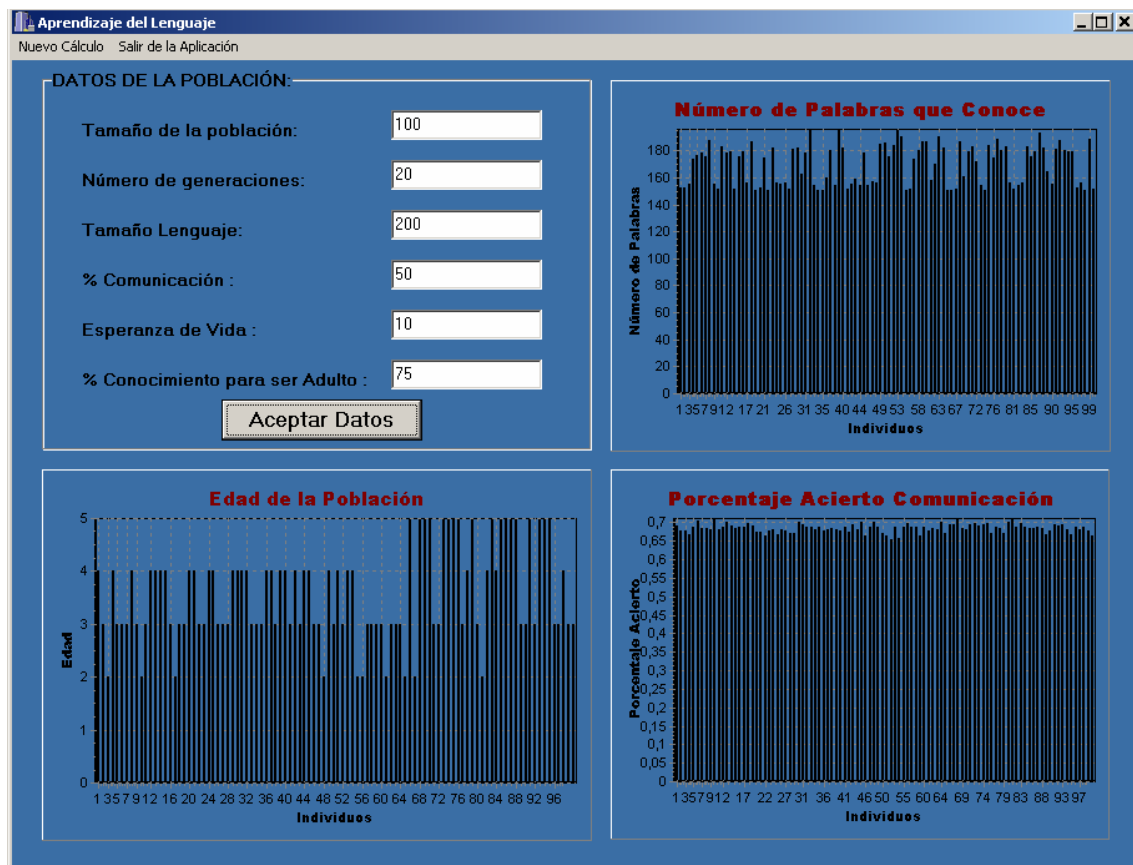
En cada paso del algoritmo se muestra:

- El número de palabras que conocen los adultos, para poder ver si los individuos de la población van conociendo cada vez más palabras o las van perdiendo.
- La edad de los adultos, para observar como evoluciona su edad.
- El número de palabras de los aprendices, comprobar como van conociendo cada vez más palabras.

- La edad de los aprendices, así podemos saber cuantas generaciones tarda en alcanzar su estado de madurez el aprendiz.

Para calcular las sucesivas generaciones se debe ir pulsando el botón **Siguiente Generación**, hasta llegar al número de generaciones que se ha especificado como parámetro de entrada.

Una vez ha terminado de ejecutarse el algoritmo evolutivo de aprendizaje, se muestran por pantalla los resultados alcanzados por la población:



Los resultados que mostramos en esta pantalla son los datos que creemos más relevantes para comprobar la eficiencia del aprendizaje.

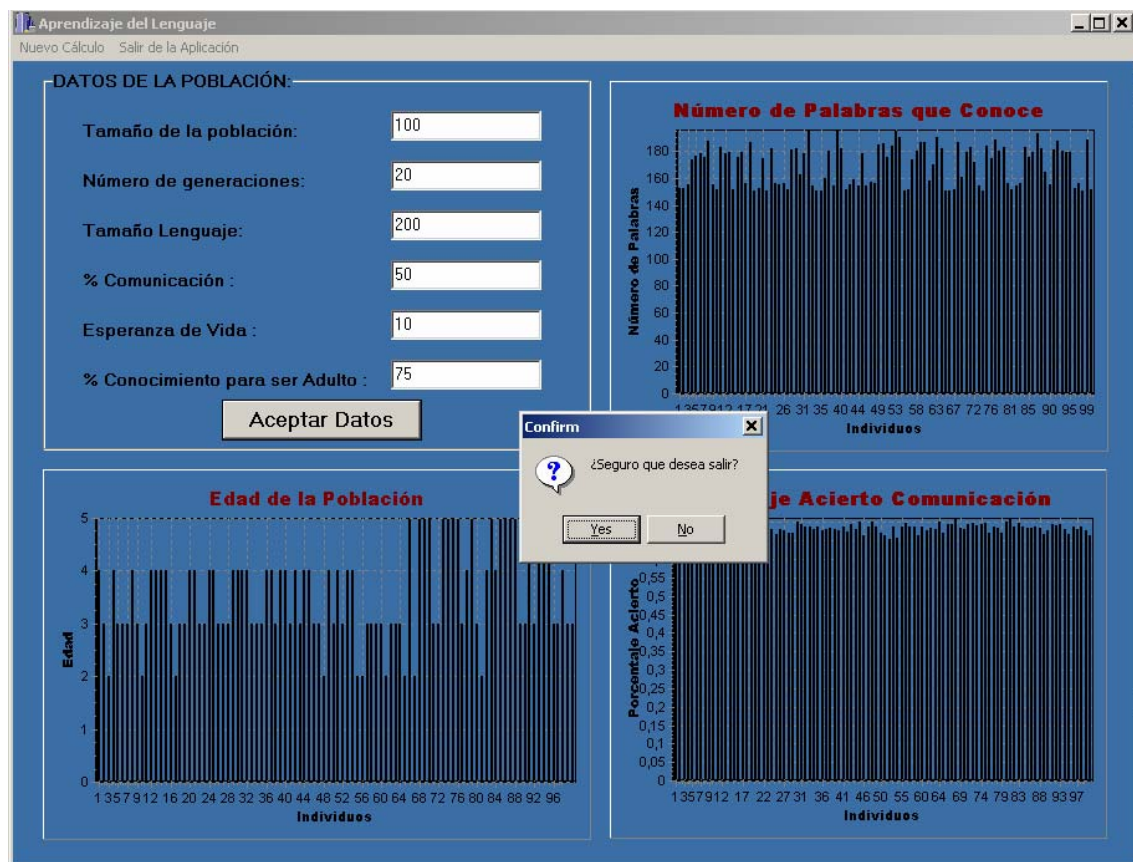
Mirando las gráficas se puede observar el número de palabras que conoce cada individuo de la población de adultos, su edad y el porcentaje de acierto medio en la comunicación con los otros individuos de la población.



Para realizar una nueva ejecución del algoritmo se debe pulsar el botón **Nuevo Cálculo**. Entonces se borrarán todos los datos de entrada y las gráficas de la anterior ejecución del algoritmo.

Para poder salir de la ejecución se debe pulsar el botón **Salir de la Aplicación**. Entonces aparecerá un cuadro de diálogo que pide la confirmación para abandonar la ejecución del programa.

Si se acepta se cierra la aplicación.



## 7.5. IMPLEMENTACIÓN DE APRENDIZAJE Y SELECCIÓN NATURAL

Para la realización de la implementación de este sistema hemos utilizado la herramienta C++ Builder.

Las estructuras de datos utilizadas han sido las siguientes:

### **Estructura de datos para el Lenguaje:**

Antes de declarar el tipo del lenguaje debemos tener el tipo de TPares, que será de lo que este formado el lenguaje.

```
typedef struct {  
    int senal;  
    int significado;  
} TPares;
```

Para el tipo lenguaje utilizamos un vector de TPares, en el que cada uno simulará una señal y su correspondiente significado.

```
typedef vector<TPares> TLenguaje;
```

### **Estructura de datos para los Individuos:**

Para el tipo individuo utilizamos un registro que constará de los siguientes campos:

El campo *genes* que será del tipo TLenguaje anteriormente definido, los campos *aptitud*, *pcomun*, *pconoc*, *pcruce* y *pmutacion* que serán números reales, y *num\_comunicacion*, *edad*, *numPalabras*, *posOriginal*, *esp\_vida* y *familia* que son campos con valores enteros y por último los campos booleanos que son reproducción, comunicación, beneficio y cruzado:

**typedef struct**

```
{  
    TLenguaje genes; //cadena de pares señal-significado  
    int num_comunicacion; //número de comunicaciones que lleva el individuo  
    int edad; //edad del individuo  
    int numPalabras; //número de palabras que conoce el individuo  
    int posOriginal; //posición original  
    int esp_vida; //esperanza de vida  
    int familia; //familia a la que pertenece el individuo  
    double pcomun; //porcentaje de comunicación  
    double aptitud; //puntuación en la comunicación con el resto de la  
        población  
    double pconoc; //porcentaje de conocimiento para ser adulto  
    double pcruce; //probabilidad de cruce  
    double pmutacion; //probabilidad de mutación  
    bool reproducción; //tipo de reproducción: true->vecinos; false->aleatoria  
    bool comunicación; //tipo de comunicación: true->vecinos; false->aleatoria  
    bool beneficio; //true mutuo, false altruista  
    bool cruzado; //indica si se ha cruzado ya en una generación determinada  
} TIndividuo;
```

### **Estructura de datos de la Población:**

Para definir el tipo de la población utilizaremos un vector de individuos:

```
typedef vector<TIndividuo> TPoblacion;
```

Las clases que forman parte de este sistema son las siguientes:

- Pantalla principal: Pide los datos de entrada necesarios para la ejecución del algoritmo y se los pasa al programa evolutivo. Cuando finaliza la ejecución muestra los resultados que hemos obtenido.
- Clase Principal: En esta clase tenemos la implementación del programa evolutivo, además del Lenguaje, la población de adultos y la población de aprendices.
- Clase Individuo: Contiene la estructura de datos de los individuos de la población.

## 7.6. MANUAL DE USUARIO DE APRENDIZAJE Y SELECCIÓN NATURAL

La pantalla que aparece al ejecutar nuestro programa es la siguiente:

**Aprendizaje y Selección Natural**

**DATOS FIJOS DE LA POBLACIÓN**

Tamaño de la Población:  Número de Generaciones:

Numero Ciclos Aprendizaje:  Tamaño Lenguaje:

**BENEFICIO EN LA COMUNICACIÓN** ☐ Mutuo ☐ Altruista

**DATOS APROXIMADOS DE LA POBLACIÓN**

Esperanza De Vida:  Porcentaje Comunicación:

Porcentaje Conocimiento:  Probabilidad Cruce:

Probabilidad Mutacion:

**Aceptar Datos** **Nuevo Cálculo** **Salir de la Aplicación**

**SOLUCIÓN:**

Porcentaje de coincidencias total:

Aptitud Media de la población:

**Aptitud** **Num comunicaciones** **Edad**

Para poder realizar la ejecución se ha de rellenar todos los datos de entrada, tanto los datos fijos de la población, como los datos aproximados.

Con los datos fijos se especifican las características de la población y con los aproximados damos los datos que se van a tomar como referencia en la creación de los algoritmos de aprendizaje de los individuos de la población inicial de adultos.

Si alguna de las casillas de los datos se deja en blanco no se podrá ejecutar el algoritmo y se mostrará un mensaje advirtiéndolo de la falta de datos.

Una vez rellenados todos los datos, para empezar la ejecución, se debe pulsar el botón **Aceptar Datos**.

Cuando termina la ejecución del algoritmo se muestran los datos obtenidos en la siguiente pantalla:

The screenshot shows a software window titled "Aprendizaje y Selección Natural". It contains several sections for configuring a simulation and displaying results.

**DATOS FIJOS DE LA POBLACIÓN**

Tamaño de la Población:	10	Número de Generaciones:	20
Numero Ciclos Aprendizaje	5	Tamaño Lenguaje	100

**BENEFICIO EN LA COMUNICACION**

☐ Mutuo ☒ Altruista

**DATOS APROXIMADOS DE LA POBLACIÓN**

Esperanza De Vida	10	Porcentaje Comunicación	50
Porcentaje Conocimiento	60	Probabilidad Cruce	0,5
Probabilidad Mutacion	0,01		

**Buttons:** Aceptar Datos, Nuevo Cálculo, Salir de la Aplicación

**SOLUCIÓN:**

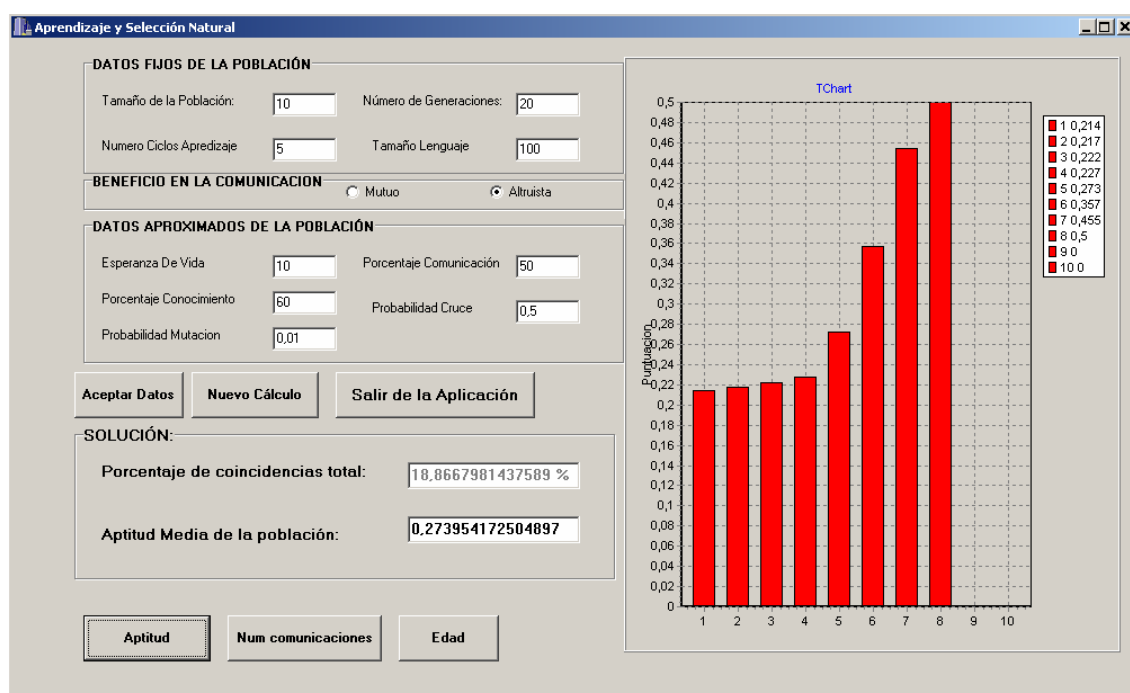
Porcentaje de coincidencias total:	18,8667981437589 %
Aptitud Media de la población:	0,273954172504897

**Buttons:** Aptitud, Num comunicaciones, Edad

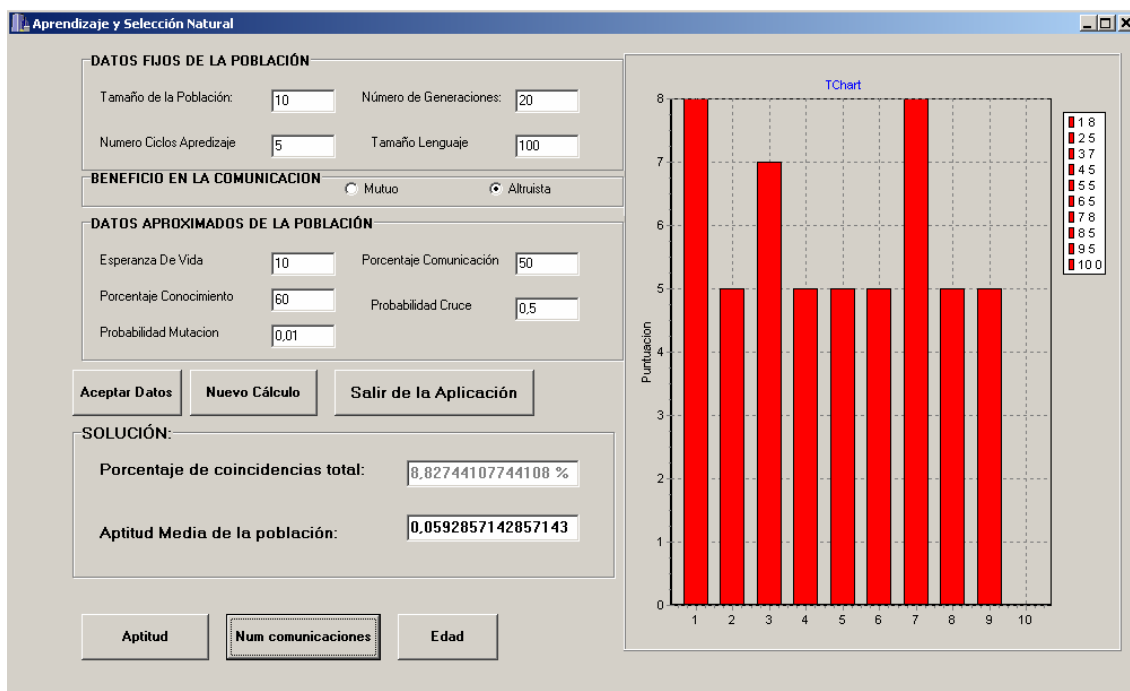
Muestra el porcentaje de coincidencia total y la aptitud media de la población.

Si se quiere observar con mayor detalle más resultados de la ejecución se puede pulsar los botones de aptitud, número de comunicaciones y edad, que mostrarán gráficamente estos valores para cada uno de los individuos de la población.

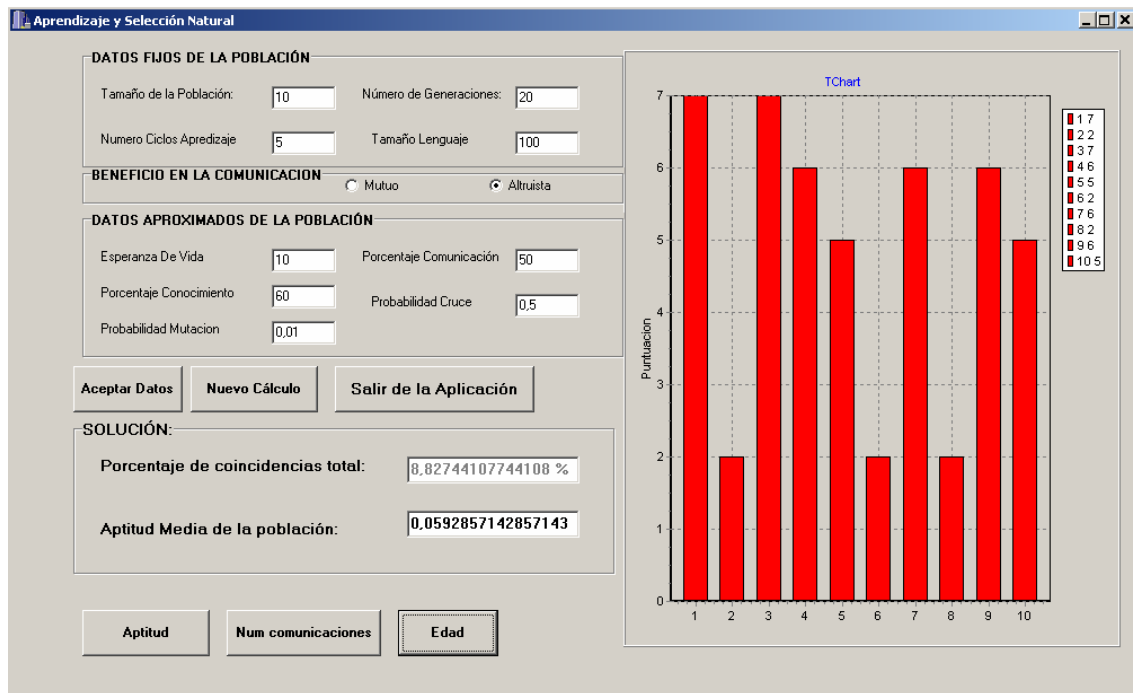
Si se pulsa el botón **Aptitud** se mostrará la siguiente gráfica en pantalla:



Si pulsamos el botón **Num comunicaciones** entonces aparecerá la siguiente gráfica:



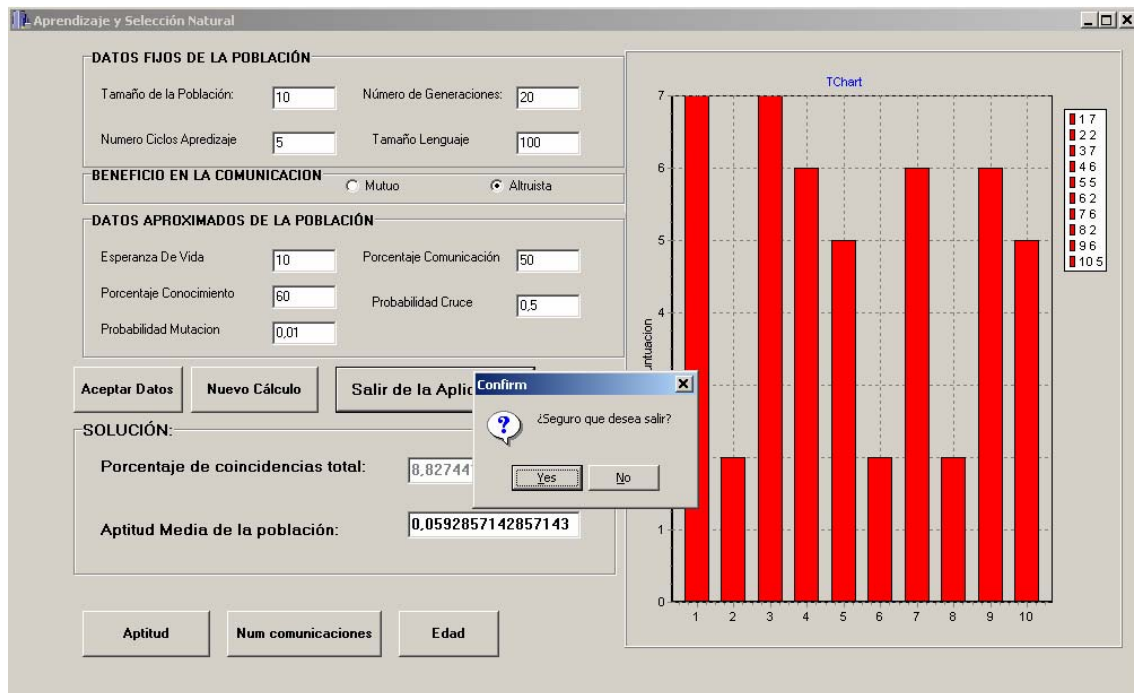
Si lo que queremos ver es la edad de cada individuo de la población, pulsamos el botón **Edad**, y mostrará gráficamente la información deseada. Se verá lo siguiente en pantalla:



Si queremos seguir calculando más resultados pulsamos el botón **Nuevo Cálculo**. Entonces se borrarán toda información de la ejecución anterior, tanto los datos de entrada, como los de salida.



Si queremos salir de la aplicación debemos pulsar el botón **Salir de la Aplicación**. A continuación aparecerá un cuadro de dialogo que nos preguntará si queremos salir. Si aceptamos, se cerrará la aplicación.



## **8. PALABRAS CLAVE**

- ALGORITMO EVOLUTIVO

30, 33, 54, 83, 86.

- APRENDIZAJE

5, 9, 47, 48, 49, 50, 52, 54, 55, 62, 67, 69, 73, 74, 79, 81, 83, 91, 97, 106, 109, 113, 116.

- COMUNICACIÓN

3, 26, 28, 35, 40, 57, 85, 88.

- INDIVIDUO

30, 32, 85, 99, 107, 113, 115.

- INNATOS

5, 8, 9, 23, 28, 30, 46.

- LENGUAJE

3, 30, 31, 33, 40, 47, 54, 55, 70, 84, 87, 99, 106, 109, 113.

- MUTACIÓN

14, 37, 91.

- POBLACIÓN

10, 12, 31, 34, 55, 56, 84, 87, 107, 114.

- REPRODUCCIÓN

10, 13, 35, 89.

- SELECCIÓN NATURAL

4, 81, 83, 97, 115, 116.

## **9. REFERENCIAS**

- [1] A. Cangelosi and D. Parisi. The emergence of a language in an evolving population of neural networks. Connection Science, 1998.
- [2] S. Kirby. Natural language from artificial life. Artificial life, 2002.
- [3] Simon Kirby. Spontaneous evolution of linguistic structure – an iterated learning model of the emergence of regularity and irregularity. IEEE Transactions on Evolutionary Computation, 2001.
- [4] Martin A. Nowak, Joshua B. Plotkin, and Vincent A. A. Jansen. The evolution of syntactic communication. NATURE, 2000.
- [5] M. Oliphant. Formal Approaches to Innate and Learned Communication: Laying the Foundation of Language. PhD thesis, Department of Cognitive Science, University of California, San Diego, 1997.
- [6] Kenny Smith. The cultural evolution of communication in a population of neural networks. Connection Science, 2002.
- [7] L. Steels. The puzzle of language evolution. Kognitionswissenschaft, 1999.
- [8] [http://es.wikipedia.org/wiki/Algoritmos\\_gen%C3%A9ticos](http://es.wikipedia.org/wiki/Algoritmos_gen%C3%A9ticos)
- [9] <http://www.sc.ehu.es/ccwbayes/docencia/mmcc/docs/t2geneticosslices.pdf>
- [10] <http://www.elrinconcito.com/articulos/Genetico/Geneticos.htm>
- [11] <http://geneura.ugr.es/~jmerelo/ie/ags.htm>
- [12] <http://www.tik.ee.ethz.ch/~zitzler/moea.html/>
- [13] <http://flanagan.ugr.es/oep/node15.html>
- [14] [http://www.ciudadredonda.org/filosofia/articulos/reflexion\\_03-06-30.pdf](http://www.ciudadredonda.org/filosofia/articulos/reflexion_03-06-30.pdf)
- [15] <http://www.lucidcafe.com/lucidcafe/library/96feb/darwin.html>
- [16] <http://www.childrenofthecode.org/interviews/daecon.htm>

